Universal Library of Multidisciplinary **Monograph**

ISSN: 3064-9986 | Volume 3, Issue 1

Open Access | PP: 40-73

DOI: https://doi.org/10.70315/uloap.ulmdi.2026.0301006

Algorithmization of the Creative Process in Commercial Photography: Methodology of Hybrid Visualization

Anastasiia Tamarina

Abstract

This monograph presents a comprehensive study of the methodology for algorithmizing the creative process in commercial photography, based on the concept of hybrid visualization, which combines neural network computational capabilities with the physicochemical parameters of analog imagery. The work systematizes the theoretical foundations of the aesthetic crisis of digital sterility, reveals the psychophysiological mechanisms of texture and color perception, and substantiates the economic necessity of transitioning from manual post-production to algorithmically controlled processing pipelines amid growing data volumes and professional burnout.

The monograph describes the technological infrastructure for training neural network profiles on paired digital RAW and film-scan datasets, where digitized film serves as a reference base for calibrating nonlinear tone mapping, color transformations, and grain modeling as a parameterized stochastic process. An engineering scheme is presented for integrating the hybrid algorithm into the production cycle, from the distribution of roles between film and digital on set to a standardized AI culling pipeline, batch neural network processing, and final human artistic validation, with a quantitative justification of efficiency via a TDABC model and calculations of labor reduction and cost savings while maintaining a premium Film Look.

Citation: Anastasiia Tamarina, "Algorithmization of the Creative Process in Commercial Photography: Methodology of Hybrid Visualization", Universal Library of Multidisciplinary, 2026; 3(1): 40-73. DOI: https://doi.org/10.70315/uloap.ulmdi.2026.0301006.

INTRODUCTION

The contemporary commercial visualization industry is located at the epicenter of a paradoxical crisis generated by the redundancy of technological capabilities. Over the past two decades, the development of digital photography has been oriented toward achieving extreme levels of sharpness, color accuracy, and noise suppression (Kirchner et al., 2021). However, upon reaching a technical plateau, when the resolving power of sensors and optics surpassed the thresholds of human visual perception, the market encountered the phenomenon of digital sterility, visual impeccability devoid of emotional resonance and markers of authenticity (Sun, 2024). In 2024–2025, this aesthetic crisis was superimposed on an acute economic imbalance: the volume of data generated in commercial shooting has reached critical levels, rendering traditional manual post-production financially impractical (Tsiavos & Kitsios, 2025).

Statistical data indicate that a typical professional photographer in the event or e-commerce segment spends from 3 to 5 hours on processing for each hour of active shooting (Guo et al., 2024). With an average volume of several thousand frames per session, the selection (culling) and color-correction stages become the bottleneck of the production cycle, absorbing most of the specialist's working time. Studies confirm a catastrophic level of professional burnout: more than half of visual content creators report symptoms of digital exhaustion due to monotonous work at a monitor (Bray et al., 2024). This situation generates a demand for a fundamental transformation of the working methodology.

In parallel with technological pressure, a global shift toward analog aesthetics (Film Look) is underway (Margadona, 2023). This trend is not a superficial form of retro nostalgia. It represents a deep psychophysiological need for organic imperfection, graininess, nonlinear color transitions, and a soft highlight gradient (Darmawan et al., 2023). In conditions where artificial intelligence can generate hyperrealistic images, the humanity of the frame, expressed through film aesthetics, becomes a key marker of quality and uniqueness.

The relevance of this work is determined by the necessity to overcome the contradiction between the market's demand for mass content production and the need to preserve its high artistic and emotional value. Algorithmization of the creative process through hybrid visualization enables automating routine operations while using objective physicochemical models of analog imagery as a reference for neural network training.

The **aim** of the monograph is to provide the theoretical substantiation and practical development of a methodology for hybrid visualization in commercial photography. This methodology aims to integrate the computational power of modern neural network algorithms with the aesthetic parameters of analog photography, enabling post-production automation while preserving the author's style.

To achieve this **aim**, the following tasks are addressed:

Investigation of the evolution of aesthetic standards and the causes of the devaluation of digital purity in contemporary visual art.

Analysis of the psychophysiological mechanisms of texture (grain) and color perception that determine the preference for analog aesthetics.

Assessment of economic losses and burnout risks within the traditional digital workflow.

Systematization of existing automation solutions and substantiation of the transition from static presets to dynamic neural profiles.

Development of an algorithm for using digitized film negatives as a training dataset for neural network calibration.

The **object** of the study is the complete production cycle of contemporary commercial photography, from the moment of light flux capture by the sensor to the delivery of the final product to the end consumer under conditions of high competition and compressed deadlines. The study examines algorithmic methods for controlling the visual characteristics of images using machine learning technologies, as well as methodological principles for synthesizing digital and analog approaches to image formation (hybrid visualization).

The **methodology** of the monograph is constructed within an interdisciplinary framework that links aesthetic theory and the philosophy of technology with the psychophysiology of visual perception and the engineering logic of machine learning. The methodological core is the concept of hybrid visualization, in which digitized film is defined as a reference (ground truth) for digital post-production calibration and as a source of physically grounded regularities in color, tonal compression, and texture that are inaccessible to subjective eye adjustment.

The practical part of the methodology is implemented by forming paired digital RAW and film-scan datasets, with procedural validation of geometric and photometric correspondence: keypoint detection, homography estimation, and subpixel registration, followed by luminance alignment in the L channel of CIELAB to exclude the influence of exposure discrepancies on color training. Neural network profile training is organized as a search for a nonlinear transformation function optimized by a combined loss function with a dominant priority on color reproduction accuracy while preserving perceptual structure and controlling texture naturalness. The robustness of the result is ensured by iterative calibration via ΔE-ITP on an independent validation set. Within the quality control loop, additional no-reference and perceptual metrics are employed to detect artifacts and maintain stylistic consistency, including monitoring typical risks in the skin tone domain using the Monk Skin Tone scale and CIELAB/ITA parameters.

The scientific **novelty** of the monograph lies in constructing a verifiable methodology of hybrid visualization, in which the film image is introduced as a measurable reference (ground truth) for digital post-production calibration. First, Film Look is formalized as a physically conditioned system of regularities of tone mapping and color formation (toe–gamma–shoulder, highlight roll-off), defined through measurable image-formation behavior. Consequently, film emulation is transferred from heuristic adjustment to a regime of controlled transformation of digital RAW. Second, a reproducible protocol is proposed for the formation and validation of digital RAW-film scan pairs (aligned paired datasets), which includes ORB/SIFT detection, RANSAC-based homography, and luminance alignment in the L channel of CIELAB, thereby minimizing spurious training signals and increasing the identifiability of learning.

Third, an applied scheme is specified for training a neural network profile focused on commercial speed, using a Neural 3D LUT and a combined loss function that prioritizes color conformity to the film reference while preserving perceptual structure and suppressing gradient artifacts. Fourth, grain is treated as a structured stochastic component of the model rather than as external noise: luminance-dependent grain synthesis and its regularization are introduced, ensuring a materially plausible texture rather than a uniform digital-noise morphology. Fifth, an iterative calibration loop (IMC) is developed, with diagnostics via ΔE-ITP and independent validation. Simultaneously, the problems of skin tone accuracy and biometric bias are operationalized through the Monk Skin Tone scale and the ITA/CIELAB metric, providing a basis for targeted correction of typical errors.

CHAPTER 1. THEORETICAL FOUNDATIONS AND THE TECHNOLOGICAL PARADOX OF CONTEMPORARY PHOTOGRAPHY

Evolution of Aesthetic Standards

The history of photographic development over the past few decades may be conceptualized as a transition from physical chemistry to computational mathematics. In the golden age of film (1930s–1980s), aesthetics were determined by the material constraints of the medium: the grain structure of silver halides, the spectral sensitivity of dyes, and the specifics of chemical development. However, with the digital revolution in the 1990s and 2000s, values such as signal purity, noiselessness, and perfect reproduction of detail became increasingly paramount (Romanov, 2024).

By 2025, the pursuit of technical perfection had reached an aesthetic dead end. Ultra-high resolution and lack of textural distortions create an uncanny valley effect in photography, where faces become overly smooth, colors become too predictable, and space is devoid of depth (Hausken, 2024). In academic discourse, this phenomenon has been designated as digital sterility (Di Natale et al., 2023).

From a neurobiological standpoint, image perception is not a passive registration of light. The human eye and brain continuously scan and interpret textures. Film grain is a stochastic (random) structure of microscopic silver crystals or dye clouds (Okada & Motoyoshi, 2021). In contrast to digital noise (a pixel grid), which is regular and induces cognitive dissonance, grain produces an effect of micro-motion.

Studies employing magnetoencephalography (MEG) and fMRI demonstrate that the brain processes color and form along distinct pathways (V1–V4 cortical areas) (Taylor & Xu, 2022). Textured surfaces (with grain) activate attentional mechanisms more effectively because they emulate the natural irregularities of the living environment. Grain functions as a kind of visual adhesive, smoothing harsh digital boundaries and generating an illusion of continuous transitions. Table 1 presents comparative characteristics of the perception of analog and digital imagery.

Table 1. Comparative characteristics of the perception of analog and digital images

Characteristic	**Analog image**	**Digital image**	**Psychophysiological effect**
Noise structure	Stochastic (random)	Regular (pixel grid)	Grain is perceived as a living texture; noise is perceived as a defect.
Color transitions	Nonlinear, subtractive	Linear, additive	Film provides a softer highlight roll-off (compression of highlights).
Dynamic range	Gentle compression in bright areas	Hard cutoff (clipping)	Film preserves detail in overexposed highlights, which is more comfortable for the eye.
Sharpness	Micro-contrast, organic	Mathematical, razor-sharp	Digital sharpness is often perceived as aggressive and flat.

The evolution of standards has led to a situation in which, by 2024, Film Look has become synonymous with premium quality. Leading brands (including high-fashion and fine jewelry segments) deliberately implement film emulation to distance themselves from mass, cheap digital content generated by smartphones (Tam & Lung, 2024).

A particularly clear example is the CHANEL N°5 campaign film See You at 5, directed by Luca Guadagnino and starring Margot Robbie and Jacob Elordi, which is explicitly described as shot on 35mm film (Azuddin, 2024). In this case, the brand's production choice is not incidental: the campaign discourse explicitly foregrounds the material attributes of film (including grain and a more sensory rendering), thereby mobilizing photochemical capture as a premium-coded aesthetic and an authenticity cue.

Fig. 1. CHANEL N°5 campaign film See You at 5 (Petrović, 2024)

In parallel, high jewelry campaigns frequently adopt an editorial image logic that privileges controlled tonal plasticity and heritage-coded visual language over maximal digital clarity. For instance, Tiffany & Co.'s high jewelry campaign featuring Abby Champion, photographed by Carlijn Jacobs, is presented as a contemporary reinterpretation of the brand's historical campaign aesthetics (Clarke, 2025). Although such sources do not necessarily specify analog capture, the campaigns' visual rhetoric aligns with film-associated perceptual markers (soft highlight transitions, chromatic restraint, and reduced clinical sharpness), thereby achieving functional differentiation from mass digital imagery.

Fig. 2. Tiffany & Co.'s jewelry campaign featuring Abby Champion (Clarke, 2025)

This confirms Walter Benjamin's theory of the aura: in an era of technical reproducibility, it is precisely controlled imperfection that becomes a signifier of authenticity and artistic value (Darmawan et al., 2023).

The concept of digital sterility is closely linked to the dematerialization of art. A digital file lacks a physical carrier. It exists as an abstract sequence of bits. In commercial photography, this has led to a decrease in the value of the moment: the ease of capturing thousands of frames devalues each individual image (Darmawan et al., 2023).

From a theoretical perspective, Heidegger's concept of Gestell suggests that contemporary digital technology enframes the world within a regime of calculative efficiency (Darmawan et al., 2023). The camera ceases to function as an instrument of poiesis (the revealing of truth) and becomes a mechanism for producing visual raw material. The return to analog aesthetics is an act of resistance to this dehumanization, an attempt to reintroduce into creative practice a process of slow seeing and material responsibility for each frame.

The Problematic Nature of the Digital Workflow

Despite the aesthetic crisis, economic realities demand unprecedented productivity from photographers. The contemporary commercial photography market in North America is estimated at 12.39 billion USD (in 2025), with projected growth to 15.84 billion by 2031 (Mordor Intelligence, 2025). This growth is driven primarily by e-commerce and social platforms, where content renewal rates are a critical factor in business survival.

The traditional workflow of a professional photographer has transformed into an economically inefficient system. Applying manual processing to data volumes of several terabytes results in significant temporal and financial losses. Given that the median hourly rate of a photographer in the United States is approximately 19.60–25.00 USD, the cost of manual post-production may exceed the cost of the shooting itself (BLS, 2023). In many segments, such as real estate photography, retoucher labor costs can amount to several multiples of revenue at the startup stage, rendering the business model unprofitable without automation (Öztaş & Arda, 2025).

Within this context, the monograph identifies two key bottlenecks that constrain the scalability of photographic businesses. Both processes pertain to repetitive operations that must be performed consistently and at scale. Consequently, they most often become bottlenecks as the number of shoots and the volume of material increase.

The first bottleneck is culling, that is, selection. In this context, it is an analytical procedure aimed at eliminating technical defects, including motion blur, closed eyes, and incorrect focus, as well as removing duplicates. For the human brain, such work proves extremely fatiguing because it requires repeated execution of homogeneous micro-decisions under conditions of limited attention. As a result, decision fatigue emerges, in which the precision of judgments decreases and the probability of missed errors or, conversely, unjustified rejection of usable frames increases (Watkins, 2024).

The second bottleneck is color correction. Its primary complexity lies not so much in finding a pleasing color solution for a single frame as in ensuring consistency across the entire series. In practice, even minor variations in shooting conditions, such as changes in the angle of light incidence or the emergence of cloud cover, lead to situations in which identical preset settings yield different results on adjacent frames. Therefore, in mass image production, the task of color correction effectively reduces to the continual leveling of the series and maintaining a stable visual standard in the face of unavoidable lighting variability (Karimipour & Witzel, 2024).

According to a survey of 188,000 photographers, the use of AI to address these tasks enabled the saving of a cumulative 89 million working hours in 2025 (Grigonis, 2025). This is equivalent to returning 12 working weeks per specialist per year. Thus, automation ceases to be a matter of choice and becomes a matter of survival.

Digital burnout in the photography industry has reached epidemic proportions. Constant exposure to emissive screens, the necessity of focusing on minute details (pixel peeping), and the monotony of movements lead to physical and psychological disorders. In 2024, 70% of creative professionals reported having experienced burnout (AMI, 2024).

It is noteworthy that attention fatigue correlates directly with the complexity of the visual material being processed. Working with clean digital files that constantly require artificial addition of contrast and texture depletes mental resources more rapidly than working with material possessing a rich internal structure (Watkins, 2024). Hybrid visualization offers a pathway to reducing this load by delegating routine operations to algorithms trained on aesthetically sophisticated exemplars.

Overview of Existing Automation Solutions

In 2024–2025, the market for photographic automation tools shifted from simple filters to complex deep-learning-based ecosystems. However, there are fundamental differences among available tools in terms of architecture and efficacy.

For a long time, the primary way to accelerate work in Adobe Lightroom was through presets. A preset is a fixed vector of metadata adjustments (for example, Exposure +0.50, Contrast +10). Its main limitation is its static nature: it does not see the content of the frame. However, alongside classic presets, it is also appropriate to consider neural network profiles. Table 2 presents a comparison of architectural approaches to processing automation.

Table 2. Comparison of architectural approaches to automation of processing

Characteristic	**Static presets (Classic Presets)**	**Neural-network profiles (AI Profiles)**
Working mechanism	Fixed parameters (XMP/DNG)	Probabilistic model based on convolutional neural networks (CNNs)
Exposure handling	Requires manual adjustment for each frame	Automatically compensates for under- and overexposure
Color processing	Global changes (HSL)	Context-aware (e.g., separating skin tones from background)
Learnability	Not possible	Profile evolves based on the user's edits
Result	Filter overlay applied on top of the photo	Individual interpretation of each RAW file

Contemporary AI solutions such as ImagenAI, Aftershoot, and Neurapix employ style transfer technology and preference-based learning. To create a personal profile, the photographer uploads 3,000–5,000 previously edited images to the cloud. The neural network analyzes thousands of parameters (from white balance to tone curves) and constructs a mathematical model of the author's style (Haim, 2026).

In 2025, three major players stood out in the market, each offering a different approach to integrating artificial intelligence into workflows. These solutions differ in architecture, data processing modality, and the degree of user control over computation, which directly influence processing speed, infrastructure requirements, and the perceived privacy risks.

ImagenAI represents a cloud ecosystem optimized for maximum processing speed, achieving up to 0.5 seconds per frame. The key advantage of the platform is the Personal AI Profile mechanism, which is trained on the work of a specific author and enables the reproduction of an individual processing style. Simultaneously, a drawback is the reliance on internet connectivity and the fact that RAW files are processed in the cloud. For some users, this raises concerns about data privacy and control over where exactly the material is processed (Haim, 2025).

Aftershoot, by contrast, emphasizes local data processing and positions Local Processing as a central characteristic of the product. This solution is convenient in situations where work involves confidential content or where access to a stable, high-speed internet connection is limited. In 2025, the platform expanded its functionality with retouching tools, including skin defect removal and tooth alignment, which, according to the source, allows users to save up to 400 hours per year. Thus, the local processing model combines enhanced control over data with a significant reduction in labor costs for typical post-processing operations (Grigonis, 2025).

Neurapix is implemented as a hybrid solution in the form of a Lightroom Classic plugin and offers the user a choice between two modes: a cloud mode that provides higher speed and a local mode that requires greater computational resources. This approach broadens use cases by allowing the process to be adapted to current hardware and network constraints. A distinct feature is the SmartPresets function, which does not require prolonged training and can be applied instantly, making the tool convenient for rapid onboarding and operational standardization of processing within the familiar Lightroom environment (Adobe Exchange, 2022).

Despite the substantial advances of AI platforms, their operation still fundamentally relies on the photographer's digital experience, that is, the aggregate of subjective decisions formed through practice in digital editors. The hybrid visualization proposed in this monograph offers a different approach and represents a qualitative leap: digitized film is used as a reference layer to calibrate these systems. In this configuration, digital processing relies not only on the author's customary aesthetic preferences but also on measurable characteristics of the film image recorded during scanning.

The necessity of this transition primarily concerns physical correctness. Film obeys a specific chemical logic of image formation, and this logic is difficult to reproduce by eye with digital editors, even for highly qualified users. The inclusion of negatives as ground truth (reference data) enables the neural network to learn the true structure of color as a regularity arising from the concrete material nature of the medium.

Furthermore, the hybrid scheme is crucial for data reconstruction. Algorithms trained on HDR film scans can more effectively restore detail in highlight regions of a digital frame because they internalize the film-specific approach to handling extreme brightness values. As a result, it becomes possible to emulate soft film clipping as a reproduction of the tone curve's stable behavior in the zone of limiting exposures.

Finally, hybrid visualization enhances the controllability of texture. However, rather than being represented as a uniform layer of digital noise over a scene, grain is typically represented as a function of the pixel's luminance/chromaticity, being governed by the physical processes underlying the effect. This means that grain cannot be represented as a layer of constant properties. Thus, in shadows, grain is coarser, while in highlights it is finer. Thus, texture can be seen as an integral part of the image-formation model and a constituent of the formation process itself.

Within the framework of hybrid visualization, the processing pipeline is described as a transformation function T from the digital space D (RAW) into the aesthetic space A (Film Look):

$$A=T(D,q)+G(L)$$

where T is a nonlinear color transformation function defined by the weights of the neural network , trained on digital–film pairs; G is a stochastic grain-generation function dependent on local luminance L.

The approach saves meaningful time, and in blind consumer trials, the product delivered is indistinguishable from a photo processed using customary techniques. Consumers are unable to tell the difference between a manually processed photo and a high-quality effect produced by an AI approach, and speed of delivery is by far the most common reason for consumer loyalty (Velásquez-Salamanca et al., 2025).

From this follows that the algorithmization of the creative process through the methodology of hybrid visualization is not merely a technical innovation but is a necessary evolutionary step that resolves the basic contradiction of economics and aesthetics in contemporary photography.

CHAPTER 2. AUTHORIAL TECHNOLOGY FOR TRAINING NEURAL NETWORKS ON THE BASIS OF ANALOG MEDIA

The transition from a purely digital workflow to a hybrid visualization methodology requires the construction of a complex technological infrastructure capable of translating the physicochemical properties of analog photoemulsion into machine-learning models. In contrast to traditional emulation methods based on subjective curve adjustments in graphic editors, the authorial technology is grounded in the use of objective data from digitized film (film scans) as ground truth, a reference standard for calibrating neural network weights (Gong et al., 2023). This enables overcoming the fundamental gap between the linear response of a semiconductor sensor and the nonlinear, subtractive color synthesis characteristic of multilayer film structures.

The novelty of the proposed approach lies in interpreting the training process as cognitive modeling of light behavior in a complex multiphase medium. This chapter elaborates on mechanisms for generating high-precision training datasets, protocols for neural network training, and systems of iterative control that maintain authorial authenticity while radically accelerating the production cycle of commercial shooting.

Formation of the Training Dataset

The effectiveness of any intelligent system correlates directly with the quality of its representative dataset. In the context of hybrid visualization, dataset formation constitutes a unique interdisciplinary challenge that requires synchronizing two fundamentally different modes of recording visual information.

A fundamental prerequisite for successful training is the creation of a set of paired images, digital RAW + film scan, captured under identical conditions. This process, known in academic discourse as the creation of aligned paired datasets, demands strict control of external variables (Mackenzie et al., 2024). To minimize parallax and ensure geometric correspondence, a specialized paired platform is used, on which a digital camera (for example, Sony Alpha 7) and an analog camera (for example, Nikon F3 or Contax 645) are mounted in the closest possible proximity, while the optical axes of the lenses are aligned with respect to the center of composition.

The shooting protocol includes capturing scenes with varying dynamic range and color temperature. It is important to emphasize that using only ideal lighting conditions leads to model overfitting and the model's inability to perform well in real commercial scenarios. Therefore, the dataset must contain:

– Scenes with hard directional light for calibrating highlight roll-off (soft compression of highlights in the shoulder of the characteristic curve).

– Complex interior compositions with light sources of different types (mixed light), enabling the algorithm to learn to interpret color shifts in the shadows.

– Portrait series with models from various ethnic groups to ensure skin tone accuracy.

Digitization of film is an act of interpretation that introduces specific characteristics into the training dataset. In the professional industry, two systems dominate, Noritsu and Fuji Frontier, each with a distinctive digital signature (Plutino et al., 2025). Table 3 presents a comparative analysis of the characteristics of the scanning systems used to form the training sample.

Table 3. Comparative analysis of the characteristics of scanning systems for forming a training sample (PhotoVision, n.d.)

Characteristic	Noritsu HS-1800	Fuji Frontier SP3000	Impact on AI training parameters
Color vector	Warm tones, emphasis on peach and pink skin tones	Cooler tones, golden skin tones, and deep cyan in the shadows	Defines the color palette the neural network will treat as the reference.
Contrast	Soft, low-contrast (RAW-like) image	High contrast, saturated colors, pronounced black point	Influences the shape of the contrast curve learned by the model.
Detail handling	Maximum preservation of detail in highlights and shadows	Tendency toward cutoff (clipping) in extreme regions	Trains the AI in dynamic-range compression methods (HDR mapping).
Grain structure	Sharp, detailed, monochromatic grain	Smoothed due to noise-reduction algorithms, chromatic grain	Forms the statistical model of texture distribution.

The table juxtaposes two typical scanner signatures, Noritsu HS-1800 and Fuji Frontier SP3000, and shows how their systematic differences propagate into the parameters the model will internalize during training. Essentially, the hardware–software scanning chain acts not as a neutral transmission channel but as a source of stable biases in color, contrast, detail handling, and noise structure. These biases establish for the neural network a norm of how skin, shadows, highlights, microtexture, and the overall tonal pattern are expected to appear.

At the level of the color vector, the devices impose different directions of color stylization. Noritsu is characterized by a warm palette with a pronounced emphasis on peach and pink skin tones, whereas Frontier tends toward a cooler rendering with a golden appearance of skin and intensified deep cyan in the shadows. The comparison between Noritsu and Frontiers is shown in Figures 3 and 4.

Fig. 3. Comparison between Noritsu and Frontiers scans (Richard Photo Lab, 2023)

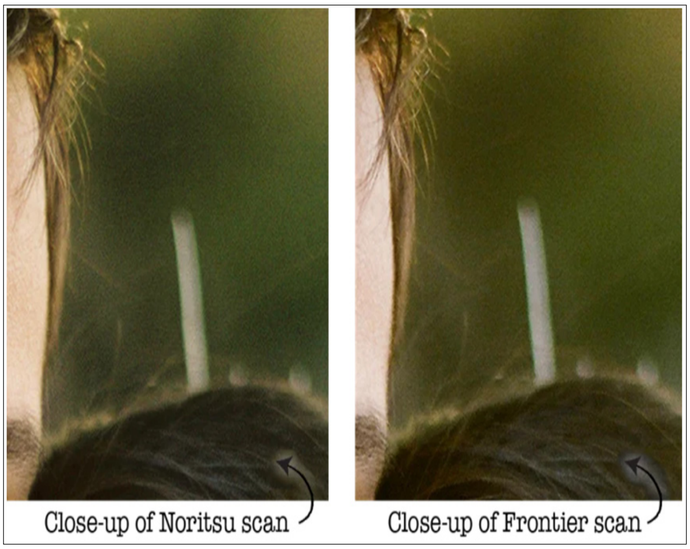

Fig. 4. Close-up comparison between Noritsu and Frontiers scans (Richard Photo Lab, 2023)

Figures 3–4 provide a concrete, side-by-side demonstration of how scanner pipelines produce repeatable perceptual differences. Figure 3 shows a clear divergence in the color vector. The Noritsu scan reads warmer, with a more pronounced peach–pink tendency in the skin tones. The Frontier scan shifts the rendering toward a cooler overall balance and a more "golden" skin impression, while the background greens appear slightly denser. The tonal feel also differs. Noritsu appears softer in the midtones, whereas Frontier looks marginally heavier in the darker regions.

Figure 4 makes the same contrast visible at the microtexture level. The Noritsu close-up preserves a grain impression that reads finer and primarily luminance-based. The Frontier close-up appears slightly smoother in high-frequency detail, with a more noticeable chromatic bias in the shadow field. Together, these paired examples demonstrate that the Noritsu/Frontier distinction is empirically observable. It manifests through a coupled shift in hue bias, tonal density, and grain handling.

For the model under training, this implies that channel distributions (and their combinations in color spaces) will nudge it toward a particular reference balance, not merely toward an averaged white balance, but toward a coherent chromatic signature. These tonal regularities may then be replicated by the network as if they were intrinsic features of the image, rather than artifacts of the individual scanner or profiles.

The next major difference is tone response. Noritsu is said to be low contrast and soft (RAW-like), whereas Frontier is high contrast, with much steeper transitions, higher saturation, and a more pronounced black point. The latter are simply the canonical curve shape, the steepness of the shadow response, where the deepest black starts, and the expected saturation variance with increased contrast. If the dataset is uniformly biased toward the Frontier type, the model will tend to generate denser, more contrasty outputs. If Noritsu predominates, the outputs will be more malleable and soft, with greater preservation of midtones.

Detail handling is related to how the system processes the dynamic range. Noritsu achieves maximal information preservation in highlights and shadows, whereas Frontier tends to clip in extreme regions. For a neural network, this is crucial: it learns not only what a photograph looks like, but also how the tails of the brightness distribution are structured statistically. Under consistent clipping, the model will regard loss of detail in overexposed or crushed areas as acceptable and may reproduce more aggressive range compression during generation or restoration. Conversely, data with superior preservation of extreme regions pushes it toward strategies of gentle roll-off and more HDR-like rendering, in which gradients in highlights and shadows are maintained over a wider interval.

Finally, grain (grain/noise) shapes for the model the statistics of microtexture. Noritsu is characterized by sharp, detailed, predominantly monochromatic grain, whereas Frontier, due to noise-reduction algorithms, smooths the structure and more often leaves a chromatic grain component. In machine learning terms, this implies different distributions of high-frequency components: their amplitudes, spectral slope, and inter-channel correlation. When trained on Noritsu-like data, the model is more likely to respect fine texture and reproduce it as neutral luminance granularity. Under a Frontier-dominated regime, it may gravitate toward stronger smoothing, encoding residual texture as colored noise artifacts. Thus, the choice of scanning source directly influences whether the model considers image cleanliness or pronounced natural graininess as the normative state.

Table 3 yields an important methodological conclusion: the scanner system acts as a latent domain factor, imprinting the dataset with predictable signatures. If the training objective is a universal model, such domain signatures must either be balanced (through source mixing) or explicitly controlled (normalization, profiling, domain adaptation).

If the objective is reproduction of a specific aesthetic, the stability of these characteristics becomes an advantage: the color vector, contrast curve, behavior at range extremes, and grain model together define the reference that the network will reproduce as the target style.

For deep learning purposes, Noritsu scans in 16-bit TIFF format are preferred, as they provide redundant information on negative density, which is critically important for training convolutional neural networks (CNNs) for detailed restoration tasks.

Consider the process of data validation and preprocessing. Even under ideal capture conditions, mechanical differences in shutters and optical designs give rise to geometric and exposure discrepancies. The dataset validation process includes a three-stage automatic alignment algorithm, illustrated in Figure 5.

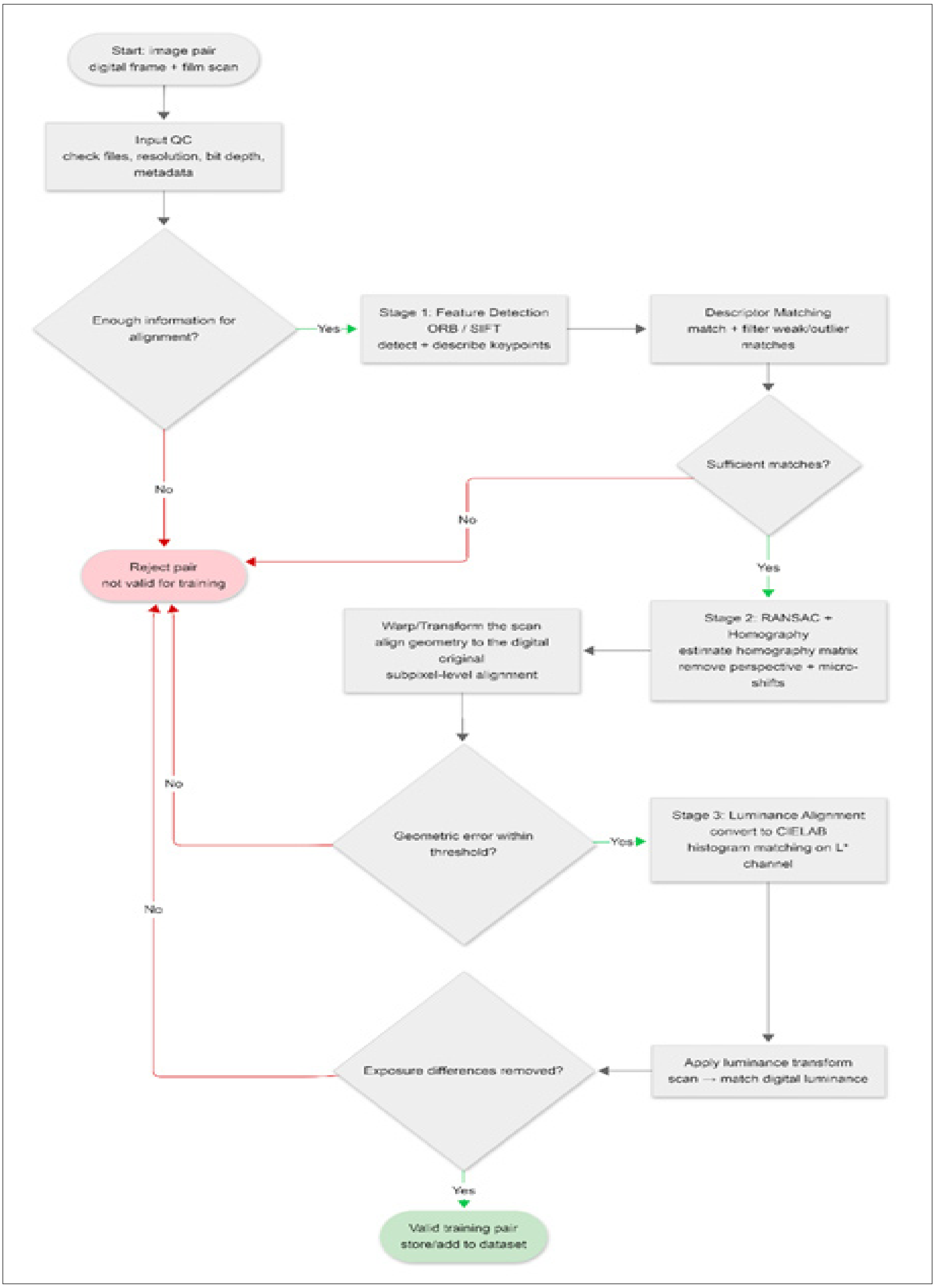

Fig. 5. Sample validation process

The first stage is feature detection using ORB or SIFT. These key points are used to match the digital frame to a film scan, which often has grain and may also include dust and scratches.

In the second phase, the Random Sample Consensus (RANSAC) algorithm estimates the homography matrix to mathematically correct the distortions and micro-shifts caused by perspective, achieving subpixel accuracy.

The third stage, the most important for color reproduction, is luminance alignment. To prevent the neural network from learning random exposure errors (e.g., due to differences in actuation speed between mechanical and electronic shutters), the luminance of the film scan is matched to that of the digital original via histogram alignment in the luminance channel of the CIELAB color space. Only after this procedure is the pair considered valid for training. Table 4 presents the protocol for geometric and brightness validation of training pairs.

Table 4. Protocol for geometric and brightness validation of training pairs

Process step	Tools / Method	Expected result
Feature detection	ORB (Oriented FAST and Rotated BRIEF)	Identification of at least 500 stable keypoints per frame.
Geometric alignment	RANSAC + Homography Estimation	Registration error (reprojection error) < 0.5 pixels.
Brightness normalization	Histogram Matching in the CIELAB L-channel	Eliminates the impact of exposure differences on color learning.
Cropping & patching	Random sampling of 256×256 or 512×512 patches	Improves training efficiency on local textures.

The protocol specifies not so much data cleaning as a formalized quality filter that reduces the entropy of the training signal by enforcing strict geometric and photometric invariants. The threshold of ≥500 ORB keypoints effectively functions as a criterion of textural informativeness: it excludes weakly structured frames in which registration is unstable and the risk of false matches increases. The combination of RANSAC and homography, along with the requirement of reprojection error < 0.5 px, elevates geometric alignment to a near-pixelwise equivalence, reducing the likelihood that the model will compensate for spatial shifts through color or texture distortion. At the same time, this constrains the domain of usable pairs: homography is adequate primarily for planar scenes or small parallaxes, so the protocol is structurally predisposed to reject cases with pronounced 3D geometry, object motion, and significant perspective differences.

The photometric component (histogram matching only in the CIELAB L-channel) deliberately suppresses exposure variations, i.e., it eliminates an easily explainable source of differences that does not relate to semantics or material properties, thereby making color learning more identifiable: the model is forced to associate color transformations with content rather than with global brightness. Simultaneously, such normalization may erase physically meaningful illumination differences (e.g., local shadows or highlights), potentially reducing the model's ability to generalize to real photometric changes. Patching with sizes 256×256 or 512×512 increases training efficiency and strengthens learning of local textures, but shifts the problem toward local statistics and may weaken control over global coherence (lighting gradients, large-scale color transitions) if these effects are not compensated by the architecture or sampling strategy. Overall, the protocol is rational as a conservative data preparation pipeline for stable training on local color–texture correspondences, but its key trade-off is a reduction of domain diversity in exchange for minimizing geometric and photometric noise.

Neural Network Profile Training Process

Training a neural network profile within the hybrid visualization framework involves finding a transformation function that minimizes the perceptual difference between the model output and the reference scan. The current technological level permits a departure from simple presets toward dynamic architectures capable of accounting for image context (Trigka & Dritsas, 2025).

To ensure high processing speed in a commercial workflow, the most effective architecture is recognized as Neural 3D LUT (neural three-dimensional lookup table) (Conde et al., 2024). In this scheme, the neural network is trained to predict optimal weights for a 3D LUT cube, which is subsequently applied to the image. This solution merges the power of deep learning with the extreme computational efficiency of traditional color correction algorithms.

The training process (Figure 6) is based on minimizing a combined loss function (Total Loss) composed of several components. This formulation implies that the model is optimized with respect to a set of mutually complementary constraints, each responsible for a specific aspect of output quality, together defining a more stable direction of learning.

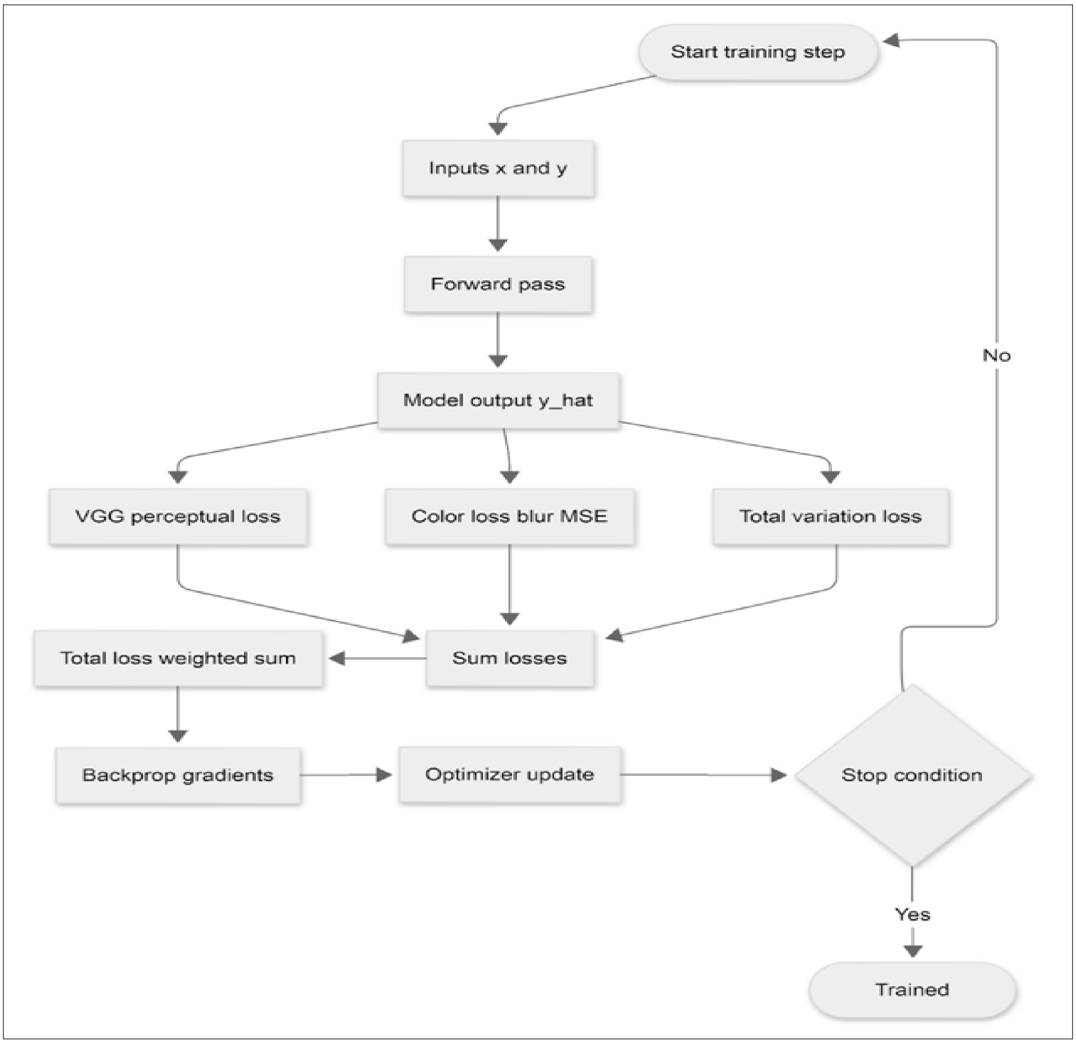

Fig. 6. AI Training Protocol

The VGG Perceptual Loss component is arranged such that, instead of directly comparing individual pixel values, the network compares features extracted by a VGG-19 model pretrained on millions of images. The criterion thus shifts emphasis from literal pixelwise coincidence to the correspondence of higher-level image structure. In this context, it is interpreted as a mechanism for preserving object structure and reducing the risk of color blotches and visual mush, which may arise when optimizing the model solely on pixel errors.

The Color Loss component is computed as the mean-squared difference between the blurred digital and film frames. The use of blur is introduced as a methodological simplification: it enables the algorithm to ignore grain and fine details, thereby separating high-frequency textural elements from the proper color information. As a consequence, the criterion focuses on the accuracy of color transitions and on white balance, i.e., on those properties that determine the coherence of the color model at the level of large tonal masses.

The Total Variation Loss (TV Loss) component is used to suppress artifacts and ensure smooth color gradients. In this context, it is described as a regularization mechanism that reduces the probability of unwanted sharp discontinuities where continuous color variation is expected. This is critically important for preventing banding, i.e., stepwise gradients, particularly noticeable in sky regions or on human skin, where the visual system is especially sensitive to disruptions of smooth transitions. Table 5 illustrates the weighting coefficients of the loss function components in the authorial protocol.

Table 5. Weighting coefficients of the loss function components in the author’s protocol

Loss function component	**Weight in total loss**	**Priority rationale**
Content Loss (L2 / MSE)	0.2	Ensures basic luminance consistency.
Perceptual Loss (VGG)	1.0	Preserves aesthetic structure and micro-contrast.
Color Similarity Loss	5.0	Prioritizes accurate film-type color reproduction.
TV-Rel (Texture Loss)	0.5	Controls organic grain distribution.

This table imposes a hierarchy of contributions of the loss-function components in the authorial protocol, with the weight distribution exhibiting a pronounced shift from photometric accuracy toward phenomenological correspondence to the visual reference. The smallest coefficient, Content Loss (L2/MSE) = 0.2, is interpreted as a deliberate weakening of pixelwise binding: this component keeps the model within a corridor of basic luminance consistency but does not dictate rigid elementwise reconstruction, which typically provokes averaging and loss of expressive high-frequency features. By contrast, Perceptual Loss (VGG) with a weight of 1.0 acts as a structural regulator, stabilizing meso-level contours and local micro-contrast at the feature-representation level. In this way, the model is optimized in a space closer to human visual sensitivity, where the alignment of feature statistics is more important than the coincidence of individual pixels.

The most dominant term is Color Similarity Loss, with a weight of 5.0, which effectively makes color reproduction (in terms of a film-type palette) the primary invariant and imposes a strong gradient priority for chromatic errors relative to textural and luminance errors. This decision can be interpreted as an assumption that the perceived authenticity of the film character is encoded more by global–local chromatic relations (tonal shifts, selective desaturation, color contours) than by the absolute accuracy of the luminance map. TV-Rel (Texture Loss) with weight 0.5 occupies an intermediate position: it does not compete with color for dominance but stabilizes the distribution of organic grain and suppresses pathological artifacts, acting as a gentle stochastic–structural regularizer. Taken together, this configuration of weights defines an optimization regime in which correct color functions as the primary criterion, perceptual structural plausibility as a secondary one, and pixel and texture discipline as supporting constraints that minimize drift and degradation of naturalness.

The key task of training is to emulate the specific nonlinearity of film. Unlike a digital file, where the relationship between light intensity and pixel value is linear, film exhibits an S-shaped characteristic curve.

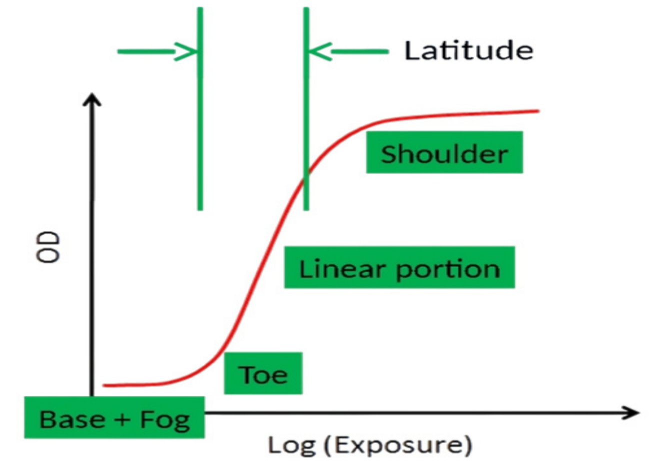

Fig. 7. S-shaped characteristic curve (Jpneylon, n.d.)

The Hurter–Driffield (H&D) curve is commonly plotted with a logarithmic scale on the x-axis, where the logarithm of relative exposure is shown. The y-axis represents the logarithm of transmission, enabling a quantitative description of how optical response varies with exposure. The low-exposure region of the H&D curve, termed the toe, extends to zero exposure. The film optical density at zero exposure corresponds to the base-plus-fog level, which reflects the inherent density of the film base together with background fog in the emulsion.

Beyond the toe, the curve enters an approximately linear region in which optical density changes nearly proportionally with the logarithm of exposure. From the standpoint of radiographic imaging, it is desirable for most of the clinically relevant signals to be recorded within this linear range, as it supports consistent tonal reproduction. At higher exposures, the curve transitions into the shoulder region, which characterizes the response under high-exposure conditions and indicates the onset of saturation in optical density.

The training algorithm is configured to reproduce three critical zones. In the shadow region (toe), film is characterized by a soft black point: detail in dark areas disappears gradually, without the abrupt cutoff typical of digital RAW. The neural network learns to imitate this behavior by retaining minimal texture in the darkest parts of the image.

In the midtone region (gamma), film produces a distinctive contrast that is perceptually volumetric. Training on scans enables the model to capture the subtractive logic of dye mixing (cyan, magenta, yellow), which differs fundamentally from the additive RGB logic of digital systems. Through this, it reproduces the characteristic interaction of tones and colors typical of film imagery.

In the highlight region (shoulder), the most complex aspect is highlight roll-off. Film can compress information in overexposed areas, helping avoid harsh white clipping. The model is trained to restore detail in bright regions of the digital frame, using HDR scans as a basis for correctly transferring structure into the high-luminance regions.

Grain, in the context of hybrid visualization, is treated as an internally organizing component of the pictorial substrate. In this sense, it is essential to distinguish digital noise, whose morphology is determined by the discrete pixel grid and the deterministic properties of the sensor and quantization, from film grain, which arises as a stochastic configuration of dye clouds or silver-based microstructures. This microphysical heterogeneity generates a quasi-random field of variation that demonstrates stable statistical invariants and is perceived by the visual system as a textural signature of the medium (Guo, Li, et al., 2024). Consequently, grain can be modeled as a structured stochastic process coupled to the semantic integrity of the image rather than as an added entropic admixture.

Within the authorial technology, dynamic grain synthesis is employed, parameterized by luminance (luminance-dependent grain), in which texture generation is governed by data and by an internalized physical law of grain manifestation across the tonal range. The neural model internalizes the dependence such that, at exposure extremes (deep shadows and high highlights), the grain component is reduced, while in midtones it reaches maximal expression. This means that grain becomes a function of local luminance and contrast: its amplitude, frequency spectrum, and probability distribution vary across the image, forming a heteroscedastic textural field. In this way, the characteristic digital uniformity of noise is eliminated, and an effect of material continuity is achieved, in which textural fluctuations are coordinated with scene photometry. 8

Figure 8 conceptualizes luminance-dependent grain synthesis as a conditional process in which local luminance and local contrast jointly parameterize the grain generator. The model first estimates photometric structure (luminance and contrast) and then maps these cues to grain statistics such as amplitude and spectral content before combining the synthesized texture with the color-transformed image. This representation formalizes grain as an intrinsic component of the imaging model. Its parameters are explicitly tied to scene photometry and remain consistent across the frame through controlled, spatially varying modulation.

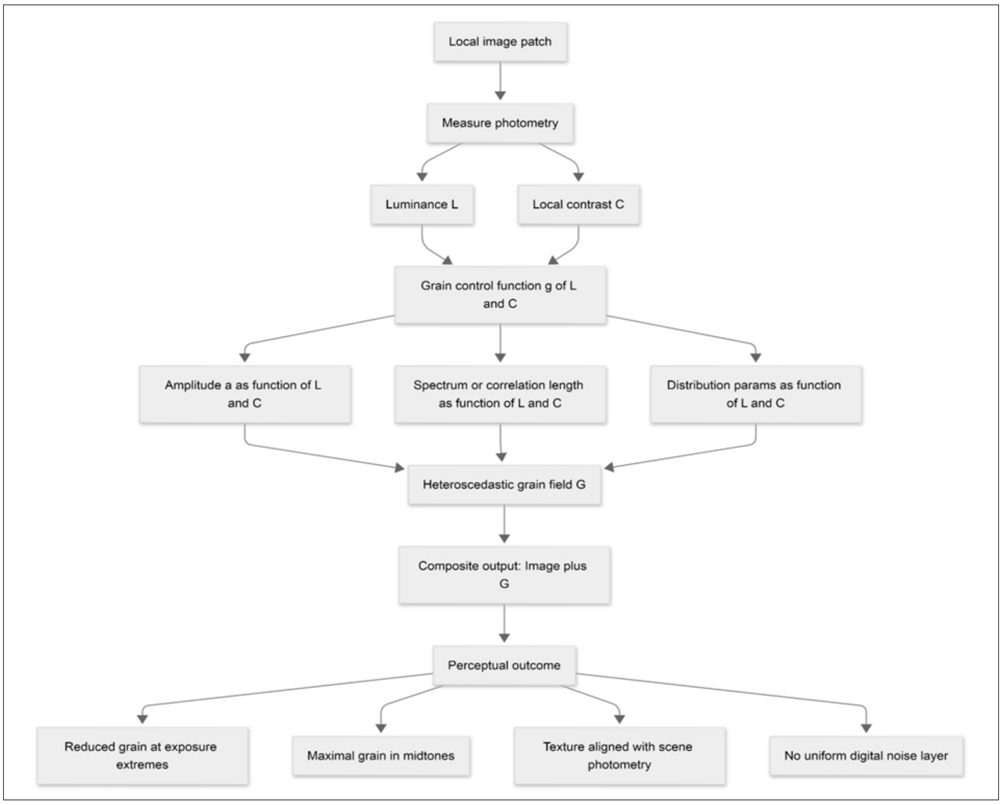

Fig. 8. Luminance- and Contrast-Conditioned Grain Synthesis Pipeline

Figure 9 depicts a unimodal relationship between relative luminance and grain strength, with maximal grain expression in midtones and attenuation toward both shadow and highlight extremes.

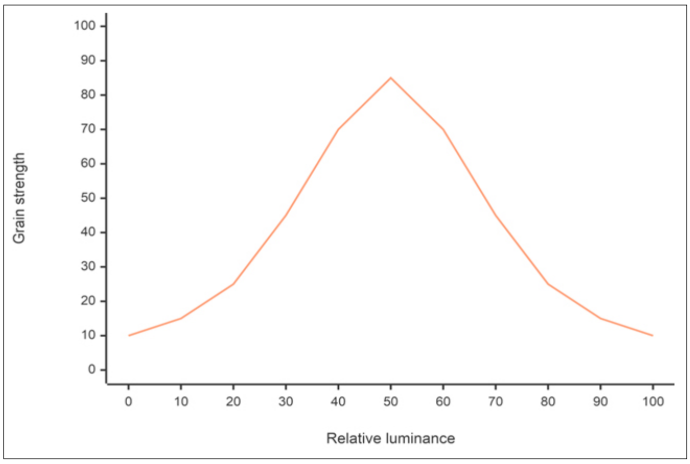

Fig. 9. Luminance-dependent grain strength (conceptual)

Such a profile is consistent with heteroscedastic texture behavior, where the magnitude of stochastic variation depends on exposure level rather than remaining constant. In practical terms, the curve motivates a grain synthesis rule that preserves perceived materiality in the tonal range where it is richest (midtones). It also limits grain in regions where excessive texture would obscure detail in shadows or disrupt smooth highlight roll-off.

Figure 10 presents grain as a heteroscedastic random field by expressing the conditional variance of grain given luminance as a function of luminance. The peak variance in the midtone range indicates that stochastic fluctuations are expected to be largest where the image carries substantial perceptual structure. Lower variance at exposure extremes constrains noise-like artifacts in visually sensitive regions. This formulation supports training and regularization strategies that target luminance-conditioned texture statistics. It enables stable perceptual continuity across the tonal scale and reduces the likelihood of uniform, sensor-like noise morphology.

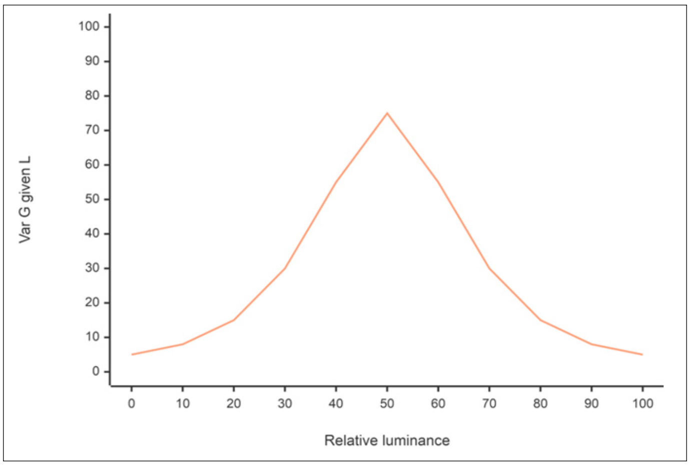

Fig. 10. Heteroscedasticity: grain variance vs luminance (conceptual)

Mathematically, this coherence is enforced by the Relative Total Variation Loss, which serves as a regularizer that compels the model to generate a textural component with controlled statistical characteristics. Optimization here aims to match grain parameters, characteristic size (correlation length), contrast (variance), and spatial–probabilistic distribution, to empirical profiles of specific film stocks. This makes it possible to interpret the synthesis as a reconstruction of a class of physical textures that differ in their stochastic signatures (for example, Kodak Portra 400 versus Tri-X 400). As a result, grain becomes both a carrier of medium identity and a mechanism of perceptual coherence, enhancing the plausibility of the hybrid image through statistically legitimate yet visually salient variability.

Iterative Calibration and Error Control

Training, Test, Correction Feedback Loop

Within the framework of this study, the training process requires iterative calibration beyond a single pass over the data, since one-time parameter optimization yields only a locally consistent approximation and leaves implicit systematic biases in color reproduction. To obtain a reproducible, metrically controlled result, it is necessary to implement Iterative Model Calibration (IMC), a procedural loop that institutionalizes demonstrable model improvement through repeated cycles of error detection, causal attribution, and targeted correction. In this formulation, training is interpreted as a sequence of controlled interventions in the joint parameter-data space, where each iteration generates new hypotheses about the dominant sources of degradation and tests them on independent data.

The key mechanism of IMC is a closed feedback loop, Training, Test, Correction, in which testing operates as a diagnostic module that produces a control signal for the next optimization step. After each training epoch, an automated error analysis is performed using the ΔE-ITP metric, which quantifies discrepancies in color characteristics between the model prediction and the reference on a perceptually relevant scale. In this way, a strict linkage is established between quality measurement and the decision to intervene: deviations are structured by patterns, distributions, and shooting conditions, thereby increasing the informativeness of errors and reducing the risk of improvements driven by random sample fluctuations.

The IMC methodology presupposes the mandatory presence of a control set of images that did not participate in the main training, so that the diagnostic stage does not degenerate into self-assessment on the same data and does not mask overfitting as progress. This set serves as an invariant arbiter of iterations: it keeps the system in a regime of external validation and renders inter-epoch comparisons substantive rather than illusory. It is important that the ΔE-ITP analysis is conducted regularly, after each epoch, to enable high-frequency monitoring of error dynamics and to identify moments when the model begins to degrade in specific lighting modes or scenes despite improvements in aggregate metrics. Thus, IMC defines not only a metric but also a measurement discipline, in which the frequency of control becomes an instrument for detecting hidden failures.

The practical implementation of IMC is described by a phase scheme of iterative film-profile calibration, where each stage has an operationally defined action and a formalized success criterion. At the initial (Alpha) phase, baseline training is performed on 1,000 pairs, and the system's state is characterized by a coarse mismatch in color differences, as evidenced by $\Delta E > 5.0$ across the overall dataset, indicating the approximation's immaturity and the presence of significant artifacts. This is followed by an analytical phase, in which hard scenes such as mixed lighting and neon sources are identified, and zones of color degradation are localized. At this point, an error cartography is produced, enabling a move from the abstract statement that the model is wrong to concrete regions of parametric and contextual risk. A subsequent corrective (Beta) stage is then implemented: fine-tuning is performed on problematic zones with a tripled weight (×3), that is, the model receives an amplified gradient signal precisely where its behavior is most unstable, and the target criterion is defined as achieving $\Delta E < 2.0$ in these zones, reflecting a transition from coarse to more subtle, perceptually acceptable calibration.

The final validation stage closes the cycle with an independent test on a separate set of 500 frames, where success is defined as $\Delta E < 1.0$ for 95% of the dataset, i.e., by the distribution and the confidence level in the stability of the results. This criterion is fundamentally important: it minimizes the risk of compensating for large errors with numerous small ones and addresses the requirement for the distribution tails, which, in practical tasks, often determine the system's perceived reliability. Collectively, IMC transforms training into a scientifically verifiable process: each iteration has a measurable input (error structure), a targeted intervention (reweighting and further training), and an independent output (validation on new data), forming a high-perplexity yet strictly causally organized loop of continuous model improvement. Table 6 illustrates the stages of the film profile's iterative calibration.

Table 6. Stages of iterative calibration of film profile

Phase	Action	Success Criterion
Initial (Alpha)	Base training on 1,000 pairs.	$\Delta E > 5.0$ on the overall dataset.
Analytical	Identify hard scenes (mixed lighting, neon).	Color degradation areas are identified.
Corrective (Beta)	Fine-tune on problematic areas with 3× weighting.	$\Delta E < 2.0$ in problematic areas.
Validation	Final test on an independent set of 500 frames.	$\Delta E < 1.0$ for 95% of the dataset.

Analysis of Typical Errors: The Problem of Skin Tone Accuracy and Biometric Bias

The analysis of typical errors in tasks involving human imagery is most appropriately begun with two interrelated classes of degradation: violations of skin tone accuracy and manifestations of biometric bias, which together undermine both the perceptual plausibility of the result and the ethical–metric soundness of the model. One of the most difficult tasks for AI remains the preservation of natural human skin appearance, since skin is not a homogeneous texture but a complex optical substrate with nonlinear dependencies among the melanin component, subsurface scattering, and the spectral structure of illumination. Studies from 2024–2025 describe the phenomenon of biometric deviation, in which algorithms trained on unbalanced datasets exhibit a systematic tendency to distort skin tones in particular ethnic groups (Melzi et al., 2024). In other words, the error assumes not a random but a directed, group-specific character. It is important to emphasize that this deviation manifests not only as a shift in mean color but also as an increase in error variance across certain phototype ranges, making the problem statistically persistent and difficult to eliminate through general quality improvement.

To transform this class of errors from heuristically described to measurably controlled, the authorial methodology employs the Monk Skin Tone (MST) scale, as shown in Figure 11, which comprises 10 types and provides a higher-resolution measurement of chromatic nuances than the classical 6-type Fitzpatrick scale (Shah et al., 2025).

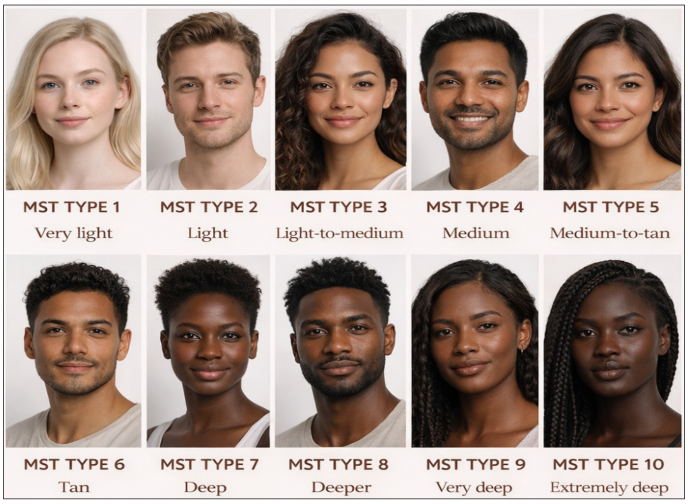

Fig. 11. Monk Skin Tone (MST) scale

The significance of this choice lies in the fact that coarse categorization of skin tones artificially collapses the multidimensional space of skin variability into an insufficient number of bins, leading the model to demonstrate acceptable aggregated metrics while still exhibiting local failures in subranges of tones. MST, by contrast, enables reconstruction of a finer error structure and mapping it to specific phototypic clusters, thereby increasing diagnostic sensitivity and reducing the risk that bias will be masked by averaging improvements. Within the logic of IMC-like control, MST functions as a discretizer of the risk space, introducing controllable granularity into the calibration procedure.

Calibration within the methodology relies on the ITA (Individual Typology Angle) indicator, which provides a geometrically interpretable, reproducible, and objectively computable criterion for linking phototypes to coordinates in the CIELAB color space. ITA is calculated on the basis of the L* and b* components, where L* expresses lightness and b* defines the blue–yellow axis; at the same time, the CIELAB system also introduces the a* component as the green–red axis, enabling the description not only of lightness deviations but also of chromatic deviations typical of machine artifacts. An important methodological implication is that AI errors cease to be interpreted as wrong color in an everyday sense: they are decomposed into orthogonal components of shift in lightness, in the blue–yellow balance, and in the green–red component, which makes causally targeted correction possible. Thus, ITA functions as a bridge between phenomenology (how skin appears) and metric description, specifying the coordinates that are shifted and by how much.

Target calibration parameters are most rationally defined in a differentiated manner across MST groups, since typical AI errors exhibit a pronounced dependence on phototype range and exposure conditions. For light phototypes (MST 1–3), corresponding to ITA > 55°, the typical tendency is toward overexposure and a shift toward yellow, which may be interpreted as a combined effect of elevated L* and a shift of b* toward positive values, amplifying the warmth of the skin. For medium phototypes (MST 4–6) with ITA in the interval 28°–55°, a loss of saturation and a gray appearance are typical, i.e., chroma degradation with relative preservation of lightness structure, indicating insufficient amplitude in the chromatic channel, especially along the a* axis. For dark phototypes (MST 7–10) with ITA < 28°, underexposure and the appearance of blue highlights are observed, manifesting as reduced L* in shadows and an undesirable shift of b* into the negative region, which intensifies a cool cast in highlight and midtone zones. This typology makes it possible to interpret bias not as an abstract unfairness but as a concrete set of directed color and tonal errors subject to engineering control.

Accordingly, correction methods are formulated as targeted transformations of CIELAB parameters aligned with the identified error profile for each group. In the Light group, corrective action is implemented by lowering L* and simultaneously adjusting a* toward red, which compensates both overexposure and yellowness by restoring the physiologically expected pinkish component. In the Medium group, it is appropriate to increase chroma in the a* channel, i.e., to restore saturation by enhancing chromatic contrast along the green–red axis, since machine models frequently collapse skin variability to neutral gray precisely in this dimension. In the Dark group, selective work with shadows is justified (raising shadow regions without destroying contours) together with balancing b* to suppress blue parasitic highlights while preserving natural depth of tone. Collectively, the described scheme defines a high-perplexity yet strictly operationalized framework: biometric deviation is fixed via MST discretization, quantitatively measured via ITA/CIELAB decomposition, and reduced through targeted corrections that minimize both visual unnaturalness and group-level systematic shifts. Table 7 presents the target calibration parameters for different skin phototypes.

Table 7. Target calibration parameters for different skin phototypes (MST Scale)

MST Type (Group)	ITA Range (degrees)	Typical AI Error	Correction Method
Light (1–3)	> 55°	Overexposure, yellow shift	Lower L*, adjust a* toward red.
Medium (4–6)	28° to 55°	Loss of saturation, gray look	Increase Chroma in the a* channel.
Dark (7–10)	< 28°	Underexposure, blue highlights	Selectively lift shadows, balance b*.

Processing Complex Lighting Schemes and HDR Scenarios

Processing complex lighting schemes and HDR scenarios in commercial photography is a fundamentally non-stationary task, as real-world setups often combine sources with differing spectral structures, for example, warm tungsten lighting and cold daylight from a window (Afifi et al., 2021). Under such conditions, what may be termed a balance conflict is formed: a model oriented toward a global estimate of the white point faces incompatible premises, because a single white point for the entire frame effectively does not exist. The paradox lies in the fact that attempts to determine a single global balance lead either to overcooling in warm zones or to yellowing in cold regions, and the resulting error becomes spatially inhomogeneous. Consequently, a correct problem formulation must proceed from the multiplicity of local illumination regimes and regard the frame as a composition of regions with distinct photometric and chromatic invariants.

In the hybrid visualization scheme, this problem is reduced via a Spatial-Aware Weighting mechanism that introduces spatial dependence into the image-processing pipeline. The neural network constructs an illumination map by estimating the distribution of luminance and, indirectly, color temperature across the frame, and then applies different weights and transformation parameters to different image segments. Thus, instead of a single global transformation, a field of local transformations is realized, where each region undergoes calibrated adjustment consistent with its lighting regime. An important methodological nuance is that the model does not simply partition the frame but is trained to align region boundaries so that transitions remain continuous and do not produce seams while maintaining local balance correctness.

The meaning of Spatial-Aware Weighting can be interpreted as a computational emulation of two complementary phenomena: local adaptation of human vision and the selectivity of film emulsion. The human visual system does not fix a single white point for an entire scene but dynamically normalizes perception based on context, allowing a neutral appearance to be maintained simultaneously across areas with different lighting conditions. Film, in turn, responds to light not only in an integral sense but also through nonlinear interactions among exposure, spectral composition, and dye density, causing its response to vary within a single frame. Translating these principles into the transformation architecture, hybrid visualization seeks not mathematically immaculate global normalization but perceptually plausible local coherence, and it is precisely this property that enhances robustness to mixed lighting.

A separate quality-control loop is devoted to preventing HDR artifacts, primarily halos around contrast edges caused by aggressive local contrast enhancement or erroneous tone-mapping reconstruction. Such halos are particularly toxic defects in a commercial context, because they are immediately perceived as traces of digital manipulation and undermine trust in the material even when color calibration is broadly correct. Therefore, iterative calibration must include not only metrics of color conformity but also criteria of structural naturalness that are sensitive to spatial distortions at boundaries. Within the described approach, this role is assigned to the BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator) metric, which enables the assessment of image naturalness without a reference original, i.e., in a mode that approximates real consumer perception.

Integration of BRISQUE into the iterative calibration loop shifts the task from the plane of physical correctness to the plane of perceptual plausibility, which is particularly critical for the final client experience. Low BRISQUE values indicate that the statistics of local structures and the spatial regularities of the image are close to natural photographic patterns and that, consequently, the result does not resemble a synthetic render or an excessively processed HDR image. In this sense, BRISQUE functions as a proxy measure of authenticity: it captures the extent to which the digital hand is visible in the microstructure of the frame. Thus, the combination of spatially aware weighting and a blind quality metric forms a two-level control system in which local chromatic adaptation is coordinated with global verification of the absence of perceptually destructive artifacts.

Thus, Chapter 2 describes the transition from theoretical concepts to rigorous engineering implementation. The authorial technology of training on analog media transforms AI from an instrument of arbitrary manipulations into a precision system that preserves the century-long experience of film aesthetics within the conditions of the contemporary digital economy.

CHAPTER 3. OPTIMIZATION OF THE PRODUCTION CYCLE: IMPLEMENTATION OF THE HYBRID ALGORITHM

Implementing the hybrid visualization methodology in the practical domain of commercial photography requires a fundamental restructuring of production processes. Whereas the previous chapters established the theoretical foundation and technological basis for training neural networks on analog media, the present chapter is devoted to the engineering and managerial integration of these solutions into the actual workflow. The contemporary ecosystem of visual production is characterized by high data entropy and strict temporal constraints, which render traditional linear methods of operation economically untenable (Silva Jasaui et al., 2024). The transition to a hybrid algorithm implies not merely the use of two types of cameras but the construction of a cyber-physical system in which the analog component is responsible for aesthetic validation and uniqueness, while the digital component, reinforced by artificial intelligence, provides scalability and speed.

Structure of Hybrid Shooting - On-site Workflow

The structure of hybrid shooting constitutes a complex logistical and creative protocol aimed at eliminating the cognitive dissonance between the slow, meditative process of analog photography and high-speed digital reportage. The effectiveness of this model depends on the strict regulation of media use, which enables it to transform the technical heterogeneity of equipment into a strategic advantage.

Task Allocation Protocol

The central element of the on-site workflow is the Task Allocation Protocol, which specifies selecting the image-capture instrument based on the scene's semantic and technical characteristics. In contrast to the intuitive approach characteristic of amateur practice, the professional hybrid algorithm is grounded in an analysis of the limiting capabilities of each medium with respect to dynamic range, color gamut, and temporal resolution.

Within the developed methodology, film (primarily 120 mm and 35 mm formats) is removed from the domain of utilitarian event recording and assigned to the creation of so-called key shots, the principal frames that construct the visual hierarchy of the project. Studies of the workflows of leading hybrid photographers show that the optimal proportion of analog material is approximately one quarter of the final selection. However, it is precisely these frames that bear the majority of the portfolio's aesthetic load (Romanov, 2024).

The use of photographic film in the workflow should be considered a strictly limited set of scenarios rather than a universal method of shooting. This is because film photography reveals its strengths most clearly in situations where the character of the emulsion's response to light, its tonal plasticity, and its image-forming peculiarities provide a predictable advantage. In other words, film is employed where the physics of the photosensitive layer functions as an expressive instrument.

In portrait photography, film is valued for its nonlinear rendering of skin tones and for the soft, gradual folding of bright areas, often described as a delicate highlight roll-off. Underlying this effect is a complex, non-uniform emulsion response to different exposure levels: as brightness increases, contrast does not break abruptly, but is redistributed smoothly. As a result, skin appears more natural and three-dimensional, with subtle transitions that, in digital capture, often have to be reproduced through labor-intensive color grading and local retouching, which can introduce secondary artifacts.

In static, low-lighted, and decorative scenes, micro-contrast and the materiality of textures also convey a sense of realism. Besides overall sharpness and frame detail, slight differences in how textures of similar luminosity (for example, fabric, wood, metal, or ceramics) capture light will greatly contribute to the realism of the whole scene. Film takes the light, and all of which is really the texture of the light, the grain pattern, local contrast, tonal compression, makes things much less ghostly and more material, and yet with a calm, non-aggressive plasticity.

Finally, film has an advantage for very high dynamic range scenes, particularly those that are backlit. It is difficult for digital cameras to fully capture highlight features without clipping. Negative film can withstand overexposure because it doesn't immediately become a white hole, and some contrast and tonality can remain for a while, so some detail can still be recovered in the scan and processing. This property is associated with the emulsion layer's construction and its ability to accumulate information in the high-luminance region, thereby creating a softer transition to white and reducing the risk of sharp, irreversible losses in bright areas.

Film frames serve as ground truth for the entire shoot. They become chromatic references to which the digital material will be aligned in post-production. This resolves the problem of digital sterility and ensures a unified perception of the series.

Within the hybrid algorithm, the digital camera serves as an instrument for operating in high-entropy, uncontrolled environments. Full-frame mirrorless systems provide the necessary data redundancy and reaction speed.

The protocol for switching to digital shooting is applied in situations where the priority is the controllability of the result and the reproducibility of technical parameters, rather than the specific tone reproduction characteristic of film. In such cases, the digital system is regarded as a more reliable measuring instrument: it provides rapid verification of exposure and sharpness, stable automation operation, and predictable file quality across a wide range of conditions. Therefore, the decision to transition to digital is a rational measure to reduce risks and losses.

The first key condition is high event dynamics, when the scene changes rapidly and requires burst sequences, precise subject tracking in focus, and minimal lag between exposures. Reportage episodes with motion, such as the bride's entrance, active dancing, or athletic elements in a commercial setup, demand burst mode and predictive autofocus, which a film camera satisfies only to a limited extent or not at all. Under these circumstances, film becomes economically costly due to material and development consumption, and technically risky, since the absence of instantaneous feedback increases the probability of missing critical moments and making errors that cannot be compensated for by repetition.

The second condition is low lighting, especially at equivalent sensitivities above ISO 3200, where film grain intensifies markedly and shifts from an aesthetic attribute to a factor of signal degradation: the discriminability of fine details decreases, shadow legibility deteriorates, and visual dirt increases in uniform areas. In these regimes, the digital sensor more often preserves a cleaner image structure owing to effective noise handling and the ability to flexibly lift shadows in RAW without such a dramatic breakdown of texture. Finally, digital shooting is used as technical insurance for critical moments: duplicating to a RAW file serves as a safety net in case of exposure or focus errors on film, enabling the event to be preserved even under local failures and reducing the overall uncertainty of the result. Table 8 illustrates the decision matrix for medium selection on the set.

Table 8. Resource allocation matrix in a hybrid workflow

Shooting Scenario	Medium	Rationale for Choice	Target Function in the Pipeline
Portrait (static)	Film (120 / 35mm)	Skin tones, micro-contrast, psychological connection	Anchor Image (color reference)
Reportage (motion)	Digital (RAW)	Autofocus, burst shooting, buffer capacity	Narrative Filler (story content)
Decor / Details	Film	Texture, color depth	Atmospheric Setter (atmosphere)
Complex mixed lighting	Digital	Ability to adjust white balance without color degradation	Technical Recovery (technical rescue)
Nighttime / low light	Digital	High ISO performance, shadow dynamic range	Low Light Documentation

The resource allocation matrix in a hybrid workflow describes targeted optimization across different types of scene uncertainty. For the static portrait and for decor/details, film is chosen as the medium with a predictable plasticity that is difficult to reproduce with digital profiles: it enhances subjectively significant parameters, skin tones, micro-contrast, the corporeality of texture, thereby increasing the probability of obtaining a reference frame. In pipeline terms, this is logical: the Anchor Image defines the chromatic and tonal beacon, while the Atmospheric Setter constructs the atmospheric layer, which can subsequently serve as a basis for stylistic consistency across the series. In other words, film here operates as an instrument for normalizing the visual language: once perception is calibrated for the viewer, it becomes easier to maintain a coherent tone for the project.

Digital shooting in the table is assigned to scenarios with high dynamics or complex light physics, those in which the cost of missing a moment or making an exposure error rises sharply. Reportage (motion) requires temporal controllability: autofocus, burst sequences, and buffer capacity reduce the risk of statistical failure, making digital RAW a rational choice for the Narrative Filler, frames that stitch the story and sustain its tempo. Under mixed lighting and at night, the priority shifts to signal recoverability: white-balance correction without substantial degradation and high ISO with acceptable detail in the shadows transform RAW into a mode of technical rescue and documentation. In general, the table articulates a principle: film serves semantics and sensation (where subtle perceptual qualities are crucial), while digital serves controllability and robustness to error (where noise, motion, and spectral ambiguity of illumination dominate).

Media Logistics and Data Management

The introduction of a physical medium into a digital production cycle generates serious logistical challenges. The risk of metadata desynchronization or the physical loss of negatives requires implementing protocols analogous to forensic chain-of-custody systems.

A labeling and synchronization protocol is needed to ensure that film scans are integrated into the digital timeline without loss and correctly matched to digitally shot frames. The critical parameter here is time: if it is recorded and propagated into the editing logic, subsequent material organization becomes not a manual reconstruction but an engineering task with verifiable reference points. Essentially, a unified temporal field is created in which any image source, regardless of medium, receives an unambiguous position relative to the course of the event.

Before shooting begins, the clocks on all digital cameras are synchronized to within 1 second so that EXIF timestamps and file system timestamps are comparable and can serve as a base scale. This reduces the entropy of subsequent assembly: frames are automatically grouped by episode, and any offsets that arise can be corrected as a simple constant shift. In practice, discipline is crucial: synchronization must be performed before the start rather than on the fly to exclude hidden misalignments that would later manifest as breaks in sequence.

For analog cameras, a Timecode Slate method emulates digital timecode using a visual marker on film. The first frame of each roll is shot so that it includes a smartphone screen showing the current time with millisecond precision (for example, via services and applications such as Time.is), or a special QR code with metadata encoded. This frame becomes an anchor: during scanning, it is converted into a measurable temporal reference, and in post-production, it can be automatically recognized, after which the frame sequence of the roll is arranged into a chronology aligned with the digital scale. In this manner, film acquires a formally defined temporal binding, and integration into the overall timeline ceases to depend on subjective memory and manual sorting. The protocol for time synchronization and integration of film scans into a digital timeline is shown in Figure 12.

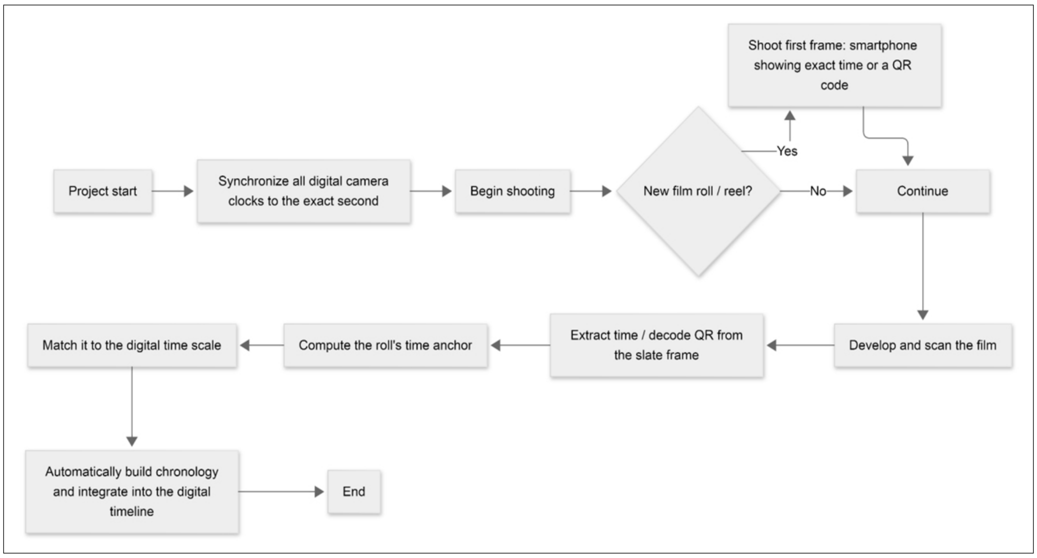

Fig. 12. Protocol for time synchronization and integration of film scans into a digital timeline

Management of physical assets in the film subsystem should be considered a system of controlled state transitions, in which each reel serves as both a material data carrier and a risk object. In contrast to a digital file, which is easily copied and automatically retains metadata, film requires an external infrastructure of identification, accounting, and protection against handling errors. Therefore, a discipline is introduced in which every instance of material has a formalized status, a traceable history, and a set of handling constraints, while deviations are recorded as rigorously as defects in a production chain.

The first control level is pre-labeling: each reel is assigned a unique ID in advance, for example, a barcode, which is then scanned into a logging application and associated with the shooting parameters. The log records the exposure index (EI), the lens used, and the lighting context, so that, after development and scanning, negative density and the emulsion's expected chromatic response can be correctly interpreted. In this way, the physical object receives a digital passport, and its behavior during the process becomes predictable: in the results analysis, material-specific characteristics can be separated from exposure errors and scene influences.

The second level is status segregation and temperature discipline, since major losses arise not at the moment of shooting but in handling between stages. Color coding of containers provides a simple yet effective mechanism for preventing double exposure and confusion: unexposed material is stored separately and is visually distinguishable from exposed material, which is placed in airtight packaging until transfer to the laboratory.

In parallel, the temperature regime is maintained: in commercial conditions, especially on location, the emulsion must remain thermally stable. Otherwise, overheating leads to irreversible color shifts and increased result variability, which is critical for high-speed films, including Kodak Portra 800. Taken together, these measures reduce the probability of loss to a level comparable to that of controlling digital material, but this is achieved not by automation but by procedural engineering. The algorithm for recording, status assignment, and thermal protection of film reels is shown in Figure 13.

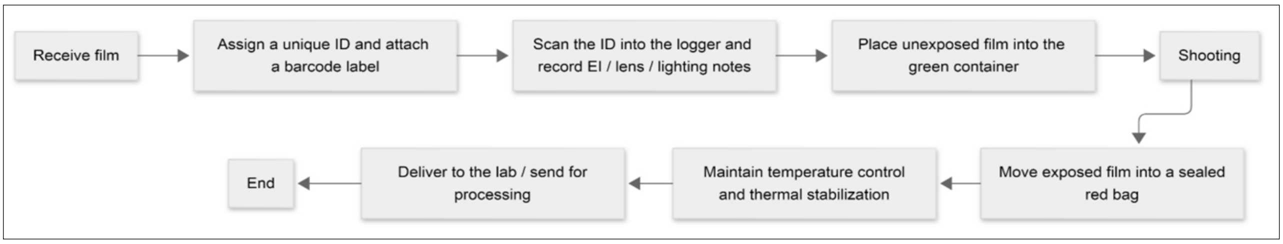

Fig. 13. Algorithm for recording, status, and thermal protection of film reels

On large commercial projects, the key figure is the DIT (Digital Imaging Technician). In the hybrid process, this role expands: in addition to backing up digital data (with mandatory checksum verification), the DIT performs on-set grading of digital frames using LUT profiles that emulate the selected film stock. This allows the art director and client to see on the monitors an image as close as possible to the final result, thereby eliminating the expectation gap between the raw digital file and the future scan.

Time Management: Accounting for the Deceleration Coefficient

The introduction of analog processes inevitably alters the shooting rhythm. The necessity of manual exposure metering (incident light), manual focusing, and film reloading requires the incorporation into the schedule of a correction coefficient.

However, this deceleration is compensated by a transformation in the structure of the photographer's attention. The limited frame resource (16 exposures per 120-type roll) shifts the work from a spray-and-pray approach to a deliberate composition. This phenomenon, described in the literature as Psychological Modulation of Workflow, leads to a paradoxical increase in efficiency (Kislinger & Kotrschal, 2021).

Time optimization within the set is achieved through cyclical switching: while the assistant reloads the film, the photographer immediately transitions to the digital camera mounted on a dual harness to capture reportage cutaways, thereby preventing downtime during the model's work.

Algorithmization of Post-production

The transition from shooting to post-processing marks a paradigm shift from manual craft to industrial automation. Post-production has customarily been done by manually linearly editing each individual file. Up to 4000–6000 frames can be generated on a shooting day, which is tedious and becomes a limiting factor in business growth. The hybrid post-production pipeline then uses neural network algorithms as the workhorse to automate repetitive tasks, while the human post-operator retains control over planned and artistic decisions.

Stage 1: Intelligent Selection - AI Culling

Primary selection of material (culling) traditionally consumes up to half of the temporal resource of post-production. In the proposed methodology, this stage is fully delegated to specialized neural networks (computer vision), integrated into software solutions such as Aftershoot, Narrative Select, or Imagen.

The AI culling process may be described as multilevel filtering, in which the input set of images passes sequentially through several independent yet coordinated evaluators. At the implementation level, this is most often an ensemble of convolutional neural networks and auxiliary models that extract features comparable to technical quality and scene semantics, and then aggregate them into a final label or rating. It is important that the system constructs a causally interpretable decision trajectory rather than sorting frames by a single scalar value: it first eliminates obvious defects, then refines suitability based on faces and gaze, and only thereafter compares close duplicates, where the differences are no longer coarse but subtle and statistical.

The first layer is typically associated with sharpness and micro-contrast detection, i.e., with assessing focus and motion traces. The algorithm constructs a saliency and sharpness map, effectively attempting to reconstruct the focal plane and identify regions that carry structural information, as measured by gradients at object boundaries. A substantial distinction of contemporary models lies in their ability to differentiate technical failure (misfocus, motion blur) from artistic softness that is part of the intended effect, for instance, bokeh or a characteristic monocle-style rendering. If the key subject, in portraiture, this is most often the eyes, falls outside the acceptable zone of sharpness, the frame is assigned a low rating or a Reject label, and the decision is grounded not in overall muddiness but in local degradation of precisely those structures that are statistically significant for narrative perception.

The next layer involves biometric analysis and gaze-direction assessment, deploying face-recognition modules and eye-state detectors. These analyze eyelid geometry, symmetry, iris position, and indirect indicators of focus in order to detect blinking, half-closed eyes, and a vacant gaze. However, the crucial aspect here is context: models are increasingly trained on scenes in which closed eyes are not an error but a semantically meaningful action, for example, in a kiss or in prayer. In such cases, a frame may be classified as intentional and not rejected, which evidences a shift from simple pattern recognition to semantic scene reading, albeit in probabilistic form.

At the final level, the system performs grouping and best-take selection, which is particularly important for series captured in burst mode. Frames are clustered by timestamps and visual similarity, after which a comparative ranking is executed within each group, a process sometimes referred to as aesthetic scoring. At this stage, the evaluation encompasses not only sharpness and eye state but also compositional balance: the distribution of masses, frame stability, centers of attention, and simplified heuristics such as the rule of thirds or approximate golden-section proportions. The outcome is the selection of one or several representative frames in each group, where the compromise between technical quality and expressiveness is optimized specifically for the given scene rather than abstractly across the entire dataset.

The result of AI culling is a reduction of the initial selection by 70–80% without human intervention. The photographer receives a pre-sorted body of material in which technical failures are hidden, and duplicates are collapsed into stacks with a designated leader. This eliminates decision fatigue and enables direct creative selection (Audiffren et al., 2023).

Stage 2: Batch AI Editing with Profile Application

The selected material (selects) proceeds to the color correction stage, where the neural profiling technology developed in Chapter 2 is applied. In contrast to static presets, which impose fixed slider offsets on all images, the neural network (for example, an architecture used in Imagen or Neurapix) analyzes the content of each frame individually.

The mechanism of contextual adaptation in the contemporary processing pipeline can be described as a probabilistic mapping f(x) → y, where the input x is a multidimensional representation of RAW data and the output y is a set of development parameters, exposure, white balance, contrast, and transformations in HSL space. A crucial nuance is that this does not constitute a single correct setting but rather a distribution of solutions: the model selects parameters that, with high probability, yield a visually coherent result under the given scene conditions. It is therefore more accurate to conceptualize the system as a statistical interpreter of light: it extracts from the raw sensor signal the structure of illumination and the semantics of the frame, and then translates these into controllable adjustment parameters.

From a practical standpoint, contextual adaptation begins with an evaluation of overall exposure and tonal range, because this is where primary ambiguity arises: the same RAW file can be interpreted as a bright, airy frame or as a dense, dramatic frame without any overt violation of physical constraints. The model resolves this ambiguity by relying on scene features: luminance distribution, presence of faces, the character of light sources, as well as statistical expectations for typical situations such as backlighting or mixed illumination. In other words, the system does not simply measure a histogram but interprets what constitutes subject, what constitutes background, and where the viewer will search for meaning, so as not to expend dynamic range on secondary regions.

A distinct critical node is white balance, where naïve gray-point correction often produces error because, in real scenes, a warm cast may be either noise or an intentional component of atmosphere. The algorithm attempts to distinguish the type of light source and its narrative role, i.e., to separate the warm hue of sunset, which is preferably preserved as a marker of time and emotional context, from the parasitic yellowish cast of tungsten lighting, which is generally perceived as contamination of skin tones and neutral surfaces. As a result, adjustment of color temperature becomes a decision as to which spectral component constitutes a meaningful signal and which constitutes interference.

After global decisions, the system progresses to local corrections via subject masking, where semantic segmentation is used to generate masks for people and background. Here, the model functions as an automated lighting operator: it may brighten faces, enhance local micro-contrast in the eye region, and simultaneously subdue an overly bright or distracting background, preserving a sense of directed lighting even in an originally flat scene. It is critical that such operations remain stable only when segmentation is accurate. Otherwise, local edits bleed into the background or clothing, creating visible seams. Consequently, in real-world systems, masks are often refined at the edges, smoothed, and adjusted to image structure so that intervention remains statistically inconspicuous yet perceptually effective.

At the level of the hybrid process, synchronization with the film reference becomes crucial because the goal is to make it appear consistent across different media. Film scans act as anchors: they define the target tonality and chromatic character, and highlight the behavior that must be reproduced in digital RAW files so that the combined gallery does not give the impression of two separate visual worlds. The neural network, trained on digital–film pairs, learns to apply a transformation that minimizes perceptual disparity with the scan across skin, shadows, highlights, and color harmony.

As a result, contextual adaptation becomes a coupled loop: the model first stabilizes the basic physical parameters of the frame, then refines color based on the meaning of illumination, then locally redistributes attention via masks, and finally aligns the result with a common stylistic denominator through film matching. This sequence is crucial because attempts to match to film before normalization of exposure and white balance lead to stylization that compensates for technical errors and amplifies artifacts. When the film reference is used as the final calibration, it functions as a metrological standard: it defines not so much a set of effects as a coordinate system in which all frames, both digital and film, become mutually compatible and visually non-contradictory. The RAW context processing algorithm with local masks and film matching is shown in Figure 14.

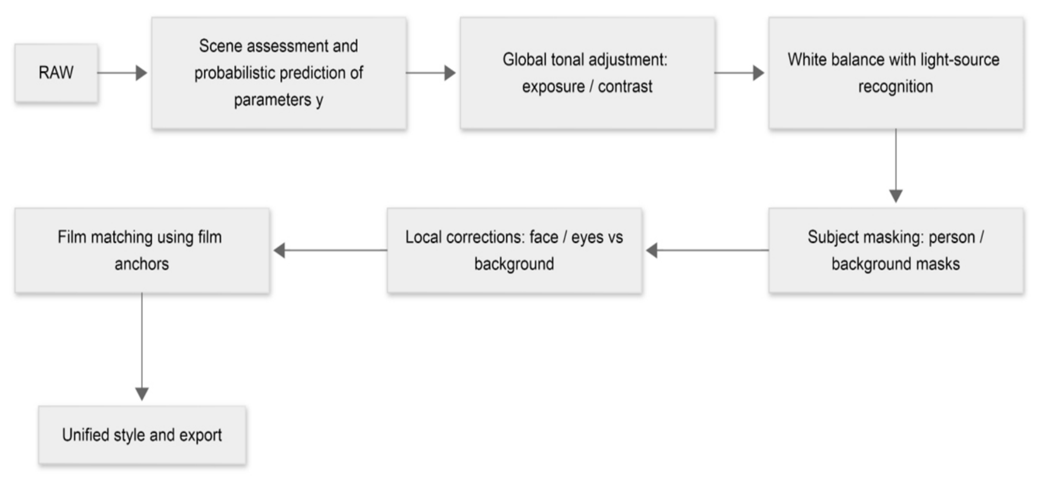

Fig. 14. RAW context processing algorithm with local masks and film matching

Processing speed is less than 1 second per image, allowing a wedding (800–1,000 photos) to be processed in 10–15 minutes of machine time, freeing tens of hours of specialist labor.

Stage 3: Final Manual Refinement - Human Touch

Even with the high accuracy of contemporary algorithms, full automation of the creative process in the premium segment remains fundamentally constrained. The reason is that the final product in this domain is evaluated by technical metrics, artistic appropriateness, ethical boundaries, and compliance with a specific client's expectations. Therefore, the Human Touch stage accounts for only 5–10% of the overall volume of operations, yet effectively serves as a stop criterion: it provides definitive artistic validation, without which automatic solutions remain probabilistic hypotheses.

In the domain of complex retouching, AI tools such as Retouch4me or Evoto can indeed automate routine steps, including frequency separation and the removal of temporary skin imperfections, such as inflammation or incidental redness. However, the boundary between a defect and an individual feature is not technical. It is cultural, ethical, and contextual. The decision on whether to retain moles, scars, expression lines, or other permanent traits is made by the retoucher, who balances the choice with norms for appropriate representation of appearance and the client's request. In this sense, automation performs a rough normalization of the surface, while the human operator is responsible for semantic finalization, ensuring that the image remains recognizable and does not devolve into a generic, impersonal face.

In a second domain, creative interpretation, human involvement is required precisely where the frame allows multiple equally legitimate readings. Under complex lighting or in tasks that require a specific mood, such as moody edits, the algorithm may select a correct but overly neutral solution because its optimal solution tends to gravitate toward statistical safety. The human operator, by contrast, may deliberately shift the balance toward drama, redistributing attention through local darkening and lightening (Dodge & Burn), as well as through cropping, which alters not pixels but rhythm and the hierarchy of meanings. Here, human intervention does not annul the work of AI but redefines it as a rough layer that can be guided along a more expressive yet still controlled trajectory.

Finally, integrity control is performed as a terminal system check for seams that are often invisible in an individual frame but become apparent when viewing the series. Viewing the gallery in Grid View is necessary to identify micro-shifts in color, density, and contrast between neighboring images, especially at the junctions of film and digital blocks, where differences in highlight character and chromatic response may generate a subtle yet perceptible dissonance. AI can evaluate frames individually and even stabilize style, but it is the human observer who detects serial inconsistency more rapidly, because the sequence is perceived as a single work in which not absolute values but relative relationships between frames are paramount. The Human Touch algorithm is shown in Figure 15.

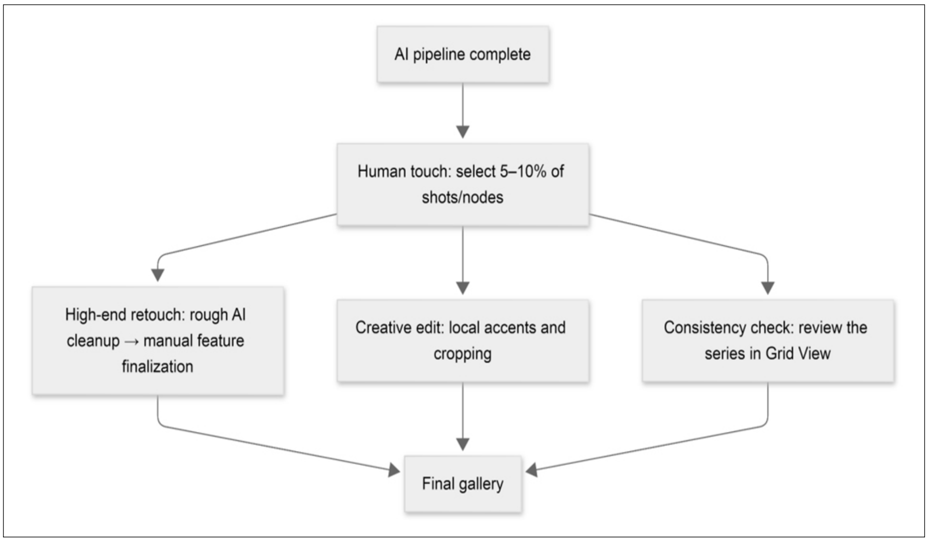

Fig. 15. Human Touch Algorithm

Below is a diagram visualizing the transformation of the production cycle.

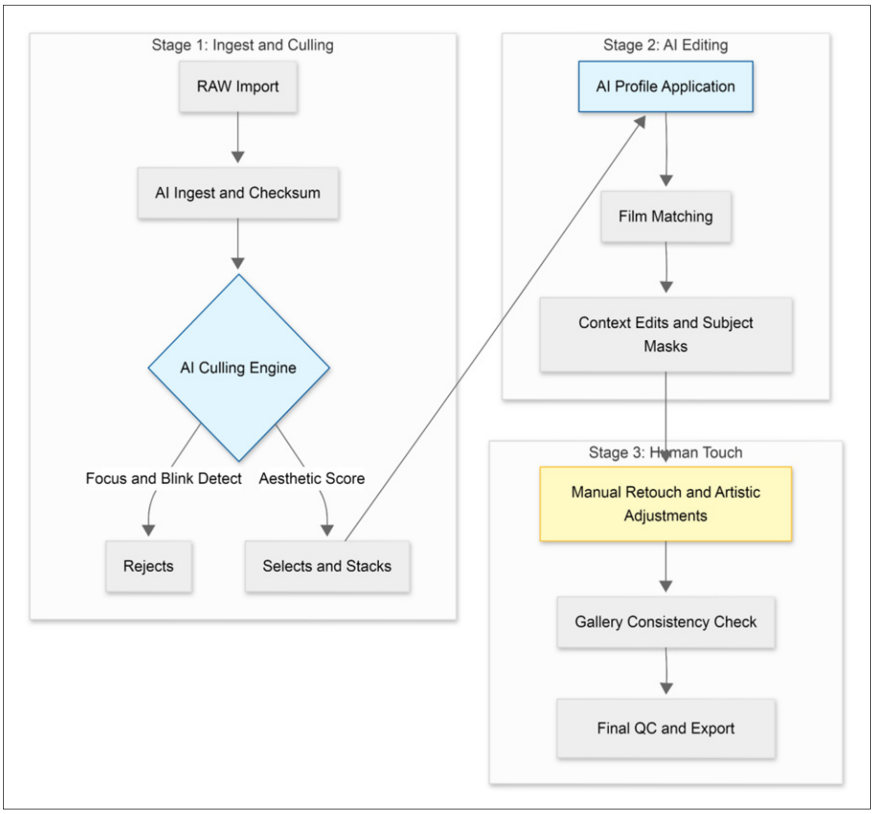

Fig. 16. Hybrid post-production pipeline with AI integration

Mathematical Model of Efficiency

The transition to a hybrid model requires rigorous economic justification. Intuitive impressions of working faster must be corroborated by quantitative metrics. For this purpose, a mathematical model of efficiency has been developed based on a time-driven activity-based costing (TDABC) method (Shakya et al., 2025). This method enables the calculation of the actual cost per unit of output (photograph) and the assessment of the return on investment (ROI) for the technologies.

To formalize the model, introduce the following variables: N_{total} - total number of source frames (RAW), $N_{processed}$ - number of processed frames, $N_{retouch}$ - number of retouched frames, N_{final} - number of final frames delivered to the client, T_{cull}^{man} - average time for manual selection (culling) of one frame, T_{edit}^{man} - average time for manual color correction of one frame, $T_{retouch}^{man}$ - average time for manual retouching of one frame, T_{setup}^{AI} - time required to configure and launch AI processes, T_{review} - time required for human verification of AI results, C_{labor} - hourly cost of specialist labor (photographer/retoucher), C_{AI} - cost of neural-network processing per frame, C_{film} - film and scanning costs (Cost of Goods Sold, COGS).

Consider the traditional model (Manual Workflow) first. In this model, all operations are performed linearly by the human operator. The total process time ($Time_{trad}$) is described by the equation:

$$Time_{trad} = \left(N_{total} \times T_{cull}^{man}\right) + \left(N_{final} \times T_{edit}^{man}\right) + \left(N_{retouch} \times T_{retouch}^{man}\right).$$

The process cost ($Cost_{trad}$) depends exclusively on labor input:

$$Cost_{trad} = Time_{trad} \text{ x } C_{labor}$$

Assume the following parameter values:

$$T_{cull}^{man} \approx 4 - 6\, sec,\ T_{edit}^{man} \approx 45 - 60 sec,\ T_{retouch}^{man} \approx 10 - 15\, min.$$

Now consider the hybrid methodology model (Algorithmic Workflow). In this model, the main labor-intensive operations (selection, basic color correction, frequency separation) are executed in parallel and autonomously by computational resources. Human involvement time is reduced to configuration and control. The total time of the hybrid process ($Time_{hybrid}$) is given by:

$$Time_{hybrid} = T_{setup}^{AI} + \left(N_{total} \times T_{review}^{cull}\right) + \left(N_{final} \times T_{review}^{edit}\right) + \left(N_{retouch} \times T_{retouch}^{hybrid}\right)$$

Here T_{review}^{cull} is the time required to confirm the AI's selection (0.5–1 s), and $T_{retouch}^{hybrid}$ is the retouch time when using AI tools (reduced to seconds). The full cost of the hybrid process includes labor, software, and analog-material expenses:

$$Cost_{hybrid} = (Time_{hybrid} \text{ x } C_{labor}) + (N_{processed} \text{ x } C_{AI}) + C_{film}$$

Next, calculate the economic efficiency (ROI) of implementing the technology. The return on investment from the transition to the hybrid model is computed using a formula adapted for service processes:

$$ROI = \frac{\left(Cost_{trad} - Cost_{hybrid}\right) - I_{setup}}{I_{setup}} \times 100\%,$$

where I_{setup} denotes the initial investment in neural network training and software acquisition.

Within the quantitative analysis, consider a typical wedding project in which the initial dataset comprises approximately 4,000 frames, the final selection is about 800 images, and 50 portraits require separate retouching. For the sake of unified calculations, specify the specialist's rate as C_{labor}= 50 USD per hour, so that labor costs are expressed as a direct product of time and rate, without accounting for overheads or seasonal coefficients.

In the traditional approach, the bulk of the time is formed by three sequential operations: primary selection, basic processing, and retouching. If we assume that culling takes 5s per frame, we would need 4000 × 5s = 20000s = 5.5 hours. If assume that the 800 final photos would take an average of 1 min (60 s) each, in total we would spend 800 × 60 s = 13.3 hours. Thus, retouching 50 different portraits at 10 minutes per image would take 50 times 10 minutes, roughly 8.3 hours.

In this cycle, the average net specialist time available is about 27 hours, or 27 x 50 = 1,350 USD in labor. In this estimate, we also factor in the cognitive overhead of manual selection and parameter tuning, in the form of attention over time and decision fatigue as the volume of material grows.

The hybrid scheme redistributes the load between automation and the human operator, thereby changing the structure of time. For AI culling and review, one may assume about 1 hour of automatic processing and about 0.5 hours of verification. For AI editing and review, approximately 0.5 hours of automated work plus about 1.5 hours of review and Human Touch can be assumed. In other words, with AI retouching, 90 images need 0.8 hours (50 × 1 minute) of time to be retouched. The total time budget is 3 hours. The labor cost is 3 × 50 = 150 USD.

In the hybrid model, only the cost component of the computation and film section is left. The AI cost is, therefore, (4000 × 0.01) + (800 × 0.05) = 80 USD. The cost of the film (approx. 15 rolls) is 300 USD. Therefore, the final cost = 150 + 80 + 300 = 530 USD. Its economic impact is 1,350 – 530 = 820 USD and the reduction in specialist working hours from 27 to 3 hours corresponds to a saving of approximately 89%.

This calculation leads to a paradoxical conclusion: even when costly analog materials are included in the cost structure, total production costs decrease due to radical automation in post-production. Moreover, the 24 hours of time freed up represent an opportunity cost: this time can be reinvested in marketing or in shooting new assignments, generating a multiplicative effect on business revenues.

Additionally, the presence of film material increases the product's perceived value, enabling a premium pricing strategy and thereby raising margins. Table 9 presents summarized efficiency indicators for adopting the hybrid algorithm.

Table 9. Comparative analysis of the efficiency of production models

Performance Metric	Traditional Model (Digital + Manual Processing)	Hybrid Model (Film/ Digital + AI)	Change / Trend
Post-production time (hrs)	27.0	3.0	-89% (Reduction)
Labor cost ($)	$1350	$150	-89% (Savings)
Variable costs (Film/AI)	$0	$380 ($300 Film + $80 AI)	Cost increase (investment in quality)
Total cost ($)	$1350	$530	-61% (Overall savings)
Aesthetic value	Standard (Digital Look)	Premium (Film Look)	Increased customer LTV
Burnout risk	High (monotony)	Low (creative control)	Improved working conditions (quality)

The table captures not only financial but also perceptual and organizational implications. For conventional digital treatment: standard aesthetic level. For hybrid treatment: premium film aesthetic. For instance, customer LTV models treat style as a premium for repeat business, referrals, and willingness to pay as customers return. In burnout models, the monotony of processing leads to the risk of burnout when using conventional digital treatments. Burnout risk with the hybrid treatment is minimized because work becomes about oversight, selection, and targeted creative innovation.

All this suggests that in a transition period between a world of handwritten texts, where the human is the ultimate arbiter and source of taste, and a world of computation, where these are mass, repeatable steps, compatible with efficiency and superior quality. Thus, the mathematical model confirms that hybrid visualization is not merely an aesthetic choice but an economically effective strategy for scaling a photography business under the conditions of Industry 4.0.

The conclusions of Chapter 3 unequivocally indicate that integrating algorithmic solutions into the structure of hybrid shooting enables overcoming the key contradiction of contemporary photography, the conflict between the requirement for high artistic quality and the demands of speed and production volume. Technologies do not replace the author but eliminate routine, returning to the photographer the primary function: creating meaning.

CHAPTER 4. PRACTICAL IMPLEMENTATION OF THE METHODOLOGY AND QUALITY STANDARDIZATION

Criteria for Evaluating Result Quality

The transition to a hybrid-visualization methodology in commercial photography requires not only technological restructuring but also the formation of a new epistemological basis for evaluating the quality of the final product. In traditional digital photography, quality criteria were often confined to technical parameters such as sharpness, noise reduction, and color accuracy (Seeram, 2023). However, within the hybrid model, where the goal is to achieve the aesthetic resonance of a film look through neural network algorithms, there is a need to quantify subjective characteristics. The problem of objectifying aesthetics is addressed through the application of the law of the Special Theory of Relativity in Image Aesthetic Assessment (SR-IAA), which postulates that aesthetic quality can be measured by determining relative preference between two comparable images within a defined time interval.

The fundamental principle of SR-IAA indicates that aesthetic evaluation is not transitive in the mathematical sense: if image A is preferred over B, and B over C, this does not guarantee that A will be preferred over C by the observer. This observation is critically important for standardizing the quality of hybrid visualization: every result produced by the algorithm must undergo cross-validation against the ground truth, which in this methodology consists of the original film scans. The quantification process aims to simulate two cognitive capacities: perception of aesthetics without a reference (by searching the input sample within the space of accumulated experience) and perception with a reference (by computing aesthetic differences between comparable pairs).

To implement this concept, a multi-attribute assessment framework (MAINet) is introduced that analyzes images through three roles: camera, photographer, and viewer. In the context of hybrid visualization, this enables formal evaluation using a set of specific metrics presented in Table 10.

Table 10. Comparative analysis of quality and aesthetics metrics in a hybrid workflow

Metric category	Tooling	Evaluation target	Impact on the production cycle
Technical quality (BIQA)	BRISQUE, NIQE, NIMA Technical	Sharpness, noise, compression artifacts, exposure errors	Automatic culling of technically defective images without human involvement.
Aesthetic value (IAA)	NIMA Aesthetic, Diffusion Aesthetics, SR-IAA (MAINet)	Composition, color harmony, and emotional response	Ranking frames by commercial appeal and style fit.

Perceptual difference	, LPIPS	Color and texture deviation from a film reference	Fine-tuning neural network weights during profile training.
Stylistic consistency	ArtBulb (DGC), CLIP features	Preserving the author's signature style across the entire series	Ensuring visual unity of the gallery (consistency check).

The table describes the hybrid pipeline as a system of separate measuring instruments, in which quality and aesthetics are decomposed into distinct categories with distinct goals. The technical BIQA metric (blind image quality assessment) answers the question 'Is the frame physically usable?', the aesthetic IAA metric answers 'Does the frame appear commercially strong?', the perceptual difference answers 'How close is the digital result to the film reference?', and stylistic consistency answers 'Is the authorial signature preserved across the series?'. This decomposition is important because the same frame may be technically flawless yet aesthetically weak, or, conversely, emotionally successful but with a critical sharpness defect. In production terms, this implies a shift from a single taste filter to a cascade of specialized filters, each optimized for its own class of errors.

The technical component relies on tools such as BRISQUE, NIQE, and NIMA Technical, which assess sharpness, noise, compression artifacts, and exposure errors without requiring a perfect reference. Their function in the cycle is to automatically reject defective frames before a human begins to invest attention in the material, thereby reducing the entropy of the input set and eliminating frames that cannot be rescued even by good processing. The next layer is aesthetic, where NIMA Aesthetic, Diffusion Aesthetics, and SR-IAA (e.g., MAINet) rank frames based on composition, color harmony, and the viewer's likely emotional response. This is a pragmatic model of commercial appeal: it aids in serial shooting by selecting frames that better align with the project's style and client expectations, especially when differences between takes are small, and the decision becomes statistical.

Perceptual difference and stylistic consistency address what usually emerges only at the final review of the series, when frames begin to disagree with one another. The EITP and LPIPS metrics measure deviations in color and texture relative to a film reference, and their practical function is calibration: they are used during training or fine-tuning of the profile so that digital processing converges toward film anchors by perceived similarity. Further, tools such as ArtBulb (DGC) and CLIP features control stylistic consistency, i.e., preservation of the authorial signature and the absence of random drifts in contrast, tone, or color between adjacent frames. As a result, the pipeline acquires a self-regulating quality: it not only rejects and ranks, but also maintains the visual integrity of the gallery, which is directly tied to premium perception and to a reduced share of manual corrections at the final stage.

A critical element of quality standardization is the use of the EITP metric. Unlike classical models, it incorporates the psychophysics of perception at high luminance levels, enabling the neural network to more accurately mimic the shoulder of the film characteristic curve (highlight roll-off) and to prevent abrupt truncation of details in highlights.

The quality-control procedure in the hybrid methodology is structured as a three-stage consistency-check protocol that progressively reduces result variability and maps the frame series into a single visual space. The logic here is close to engineering practice: first, random fluctuations that carry no meaning and impede comparison are removed. Next, frames with probabilistic defects are filtered out. Only after that is the human introduced as the final arbiter of artistic suitability. This order is important because subjective evaluation becomes more stable when the input material has already been cleansed of technical noise and does not provoke misleading impressions.

At the first stage, automatic alignment of luminance and color in CIELAB space is performed, which is convenient in that it brings numerical operations closer to human perception of differences. The system aligns the series to a common base density and neutral color, minimizing the impact of minor exposure deviations and random color shifts between frames. Conceptually, this is calibration: it makes neighboring images comparable so that subsequent decisions are not distorted by the trivial instability of initial parameters.

At Stage 2, smart selection is enabled, in which an AI culling algorithm based on convolutional neural networks analyzes frames using a large set of features. It works by detecting the micro-blur, focus misses, and the semantic importance of the defect (whether it is on a major perceptual parameter such as the eyes, their expression, the gesture, or the interaction). Thus, the presence of closed eyes, an involuntary distortion of the face, or an indifferent gaze will be penalized more than a softness in a secondary background. As a result, the automatic layer functions as a probabilistic filter, reducing the volume of material to a manageable set and increasing the density of successful frames in the series.

At the third stage, final validation by a human is performed, and its role is fundamentally not reducible to formal checking. The photographer confirms whether the result corresponds to the client's aesthetic expectations, relying on a subjective sense of reality of the texture, that is, on how natural the skin, light, grain, and tonal transitions appear within the context of the entire gallery. Here, the limits of automation are evident: algorithms can bring the series close to a consistent state, but the decision about where the boundary lies between technical correctness and artistic plausibility remains contextual and depends on taste, task, and agreement with the client.

A special place in the evaluation system is occupied by blind tests. Studies show that when high-quality neural profiles trained on paired digital–film datasets are used, respondents (including professional critics) are, in most cases, unable to distinguish algorithmic emulation from the physical medium (Osińska et al., 2025). This confirms the hypothesis that aesthetic experience can be fully algorithmized, provided that a physically grounded training base is preserved. To visualize the quality verification process within the hybrid methodology, a schematic of the control pipeline is presented below.

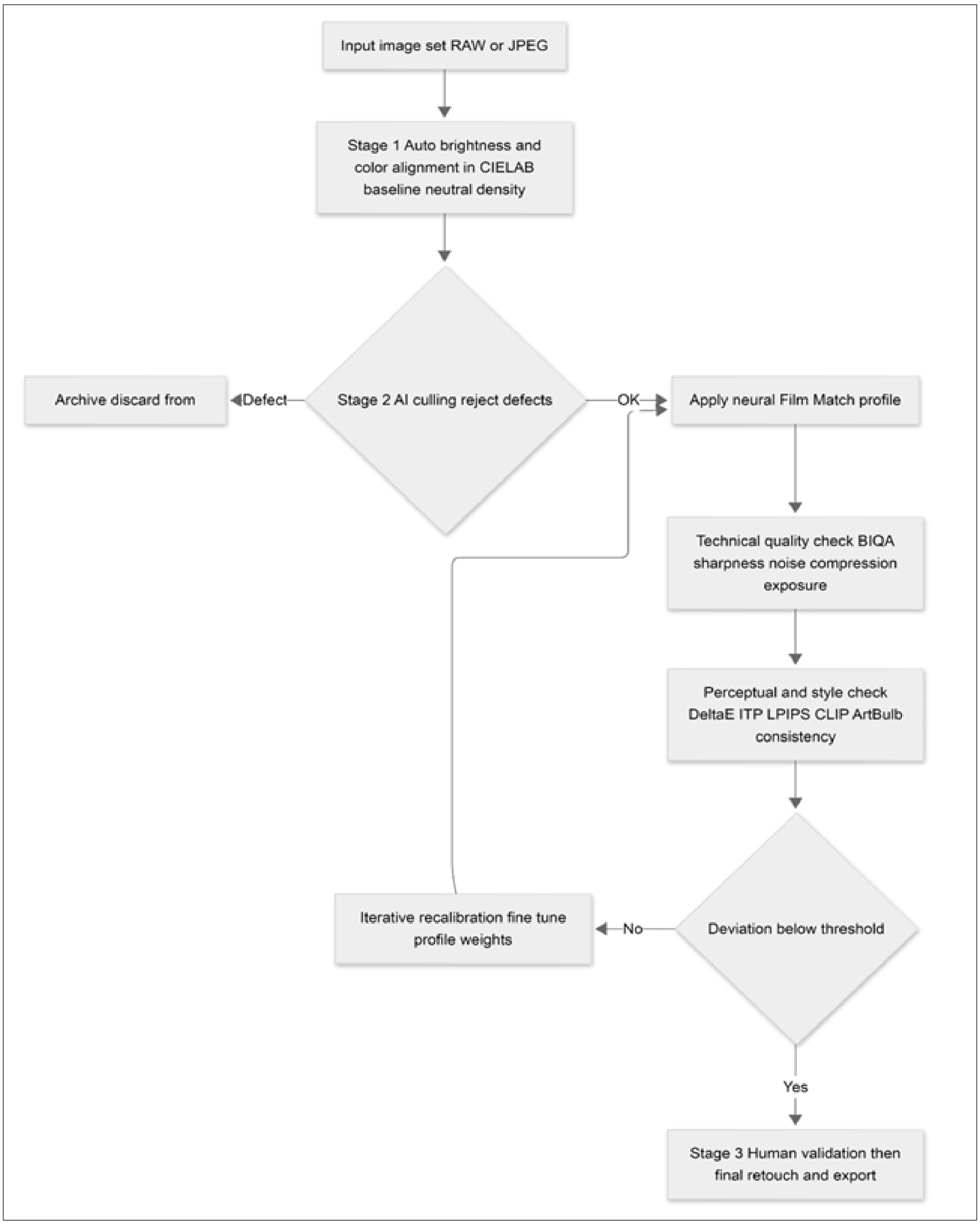

Fig. 17. Quality control pipeline diagram in hybrid imaging

Scalability of the Methodology

The scalability of the proposed hybrid visualization methodology is determined by its ability to integrate into production cycles of varying complexity, from individual freelance photographers to global media agencies. The introduction of AI into creative industries has moved from the experimental phase to full-scale deployment (Bervar et al., 2026). Almost 90% of companies regularly use AI in at least one business function, indicating a stable increase in AI adoption within organizations (Singla et al., 2025).

The methodology of hybrid visualization demonstrates high adaptability in three key dimensions: genre, technology, and economics. In the genre dimension, the approach scales successfully to wedding, reportage, fashion photography, and e-commerce. Table 11 demonstrates the specifics of scaling the methodology across genres of commercial photography.

Table 11. Specifics of scaling the methodology by genres of commercial photography

Photography genre	Data volume (shots per project)	Key automation agent	Scaling outcome
Wedding reportage	3,000–6,000	AI Culling + Personal AI Profile	Delivery time reduced from 2 months to 48 hours; 89 hours saved per project.
E-commerce (Catalog)	10,000+	Batch Image Generation + Scene Swaps	Cost per image reduced by 90%; ability to process 1,000 SKUs overnight.
Fashion / Editorial	200–500	High-End AI Retouching + Film Anchor Match	Higher perceived value thanks to a distinctive film look.
Real estate	50–100	AI HDR Mapping + Perspective Correction	Removes digital sterility in interiors; listing conversion up 40%.

The table shows that the hybrid methodology does not scale as a one-size-fits-all approach but rather by tuning the dominant automation agent to the typical structure of the genre. Genres differ primarily in data volume per project and in the nature of the key risk: in wedding reportage, the risk is missing a successful moment among thousands of takes; in e-commerce, becoming overwhelmed by standardization and repetitive scenes; in fashion/editorial, losing premiumness due to an overly digital texture; and in real estate, producing an interior that is technically correct yet visually sterile. Accordingly, the result metrics differ: in some cases, the main outcome is delivery speed; in others, cost per frame; perceptual value; and in yet others, listing conversion, i.e., a behavioral metric of the audience.

In high-volume scenarios, the table captures the most industrial effects. For wedding reportage with 3,000–6,000 frames, the key module is the AI culling and personal AI profile bundle, which transfers selection and basic stylization from manual mode to controlled review mode. The reported effect is a reduction in delivery time to 48 hours and savings of about 89 hours per project. In e-commerce, with 10,000+ images per project, the main agent is batch generation and scene swaps, reflecting conveyor logic: uniformity and throughput are paramount. Here, the outcome is formulated as a 90% reduction in per-image cost and the ability to process around 1,000 SKUs overnight, i.e., productivity increases not linearly but in leaps, because it is precisely the repetitive component of labor that is automated.

In low- and medium-volume genres, efficiency manifests less in time than in perception quality and the commercial effect beyond the image. In fashion/editorial, with 200–500 frames per project, the leading roles belong to high-end AI retouching and film anchor match, i.e., calibration of digital material to a film reference: this is not acceleration for its own sake but management of a recognizable style that raises the perceived value of the series. In real estate, with 50–100 frames, the key agent is AI HDR mapping and perspective correction. Here, the stated result is a 40% increase in listing conversion due to the elimination of digital sterility and a more natural tonal rendering of the interior. Taken together, the table describes the hybrid methodology as an adaptive set of procedures in which automation is selected according to data type and genre-specific target metrics: speed, cost, premiumness, or conversion.

Technological scalability within the hybrid methodology is grounded in the principle of Hybrid by Design, that is, in designing the process from the outset such that the digital and analog components are not stitched together post factum but are compatible at the level of data architecture and operations. In this sense, scalability is understood as the system's ability to expand without exponential growth in manual labor and without loss of controllability. Conceptually, this implies vertical integration, in which individual stages are linked into an end-to-end pipeline with shared control rules, and horizontal integration, in which new tools and modules are added to the system without disrupting the core logic (Shi & Shen, 2025).

In 2025–2026, this logic is expressed in the transition from static plug-ins, which operate locally and require manual coordination, to cloud and agent infrastructures, in which operations become services and control becomes scenario-based and event-driven. Practically, this means that processing ceases to be a set of buttons within a single application and becomes a distributed process in which project state, metadata, and control points exist independently of a particular workstation. Such a shift increases reproducibility of results and reduces the likelihood of silent errors, because each stage leaves a formal trace: what was done, with what tool, and under which decision rules.

AI agents in the contemporary interpretation are regarded as autonomous systems capable of planning actions, executing multi-step procedures, and adapting to context without continuous human supervision (Mayer et al., 2025). For a large photo studio, this implies the emergence of a specialized ecosystem of agents, where functions are not mixed but distributed by role: one agent is responsible for import, project structure, and checksum verification; a second for stylistic processing based on a neural profile and anchored references; and a third for assembling derivative products such as social media previews, contact sheets, or automated photo book layouts. Under such role separation, the system resembles a laboratory protocol: each module has a limited domain of responsibility, and the overall result is achieved through consistency of interfaces and data-transfer rules.

Economic scalability in this framework is expressed not only in direct savings but also in a more complex dynamic of return on investment. According to estimates, applying hybrid principles can provide approximately a threefold ROI advantage over traditional methods over a five-year horizon (O'Brien et al., 2025), and this effect is explained by a dual mechanism. On the one hand, costs of manual operations, including retoucher labor and time spent on routine sorting, are reduced. On the other hand, opportunity costs decrease: the time freed for the photographer or producer can be redirected toward marketing, sales, and expansion of the partner network. As a result, the economic gain becomes multiplicative: the system not only does the same thing more cheaply, but also creates additional growth opportunities by accelerating the cycle and increasing product predictability.

However, the scaling process confronts a skills deficit barrier. IThe top barrier for adopting AI into workflows in 2025 was a lack of skill, and the industry is responding with efforts to raise AI fluency (Babashahi et al. 2024). The hybrid visualization methodology proposed in this work overcomes this barrier through its intuitive interface, so the photographer does not need to be a programmer. They only need to provide a high-quality training dataset consisting of their best works and film scans.

Prospects for scaling up to 2030 are linked to the emergence of Sovereign AI, in which studios and agencies deploy their own models within protected environments, ensuring strategic independence and security of client data. This will enable scaling not only in volume but also in quality of service personalization, creating digital twins of the creative style of each leading author within the studio.

Ethical and Legal Aspects

The development of algorithmization of the creative process in photography foregrounds a number of acute legal and ethical issues, which by 2025 have become the subject of detailed regulation. The central problem is the legal status of a neural profile and its products. At the core of the discussion lies the idea/expression distinction: copyright protects concrete expressions but not general ideas, styles, methods, or techniques (Drassinower, 2022).

According to the U.S. Copyright Office (USCO) report of January 2025, works created solely by AI, without the guiding hand of a human, are not subject to copyright protection (U.S. Copyright Office, 2025). However, the methodology of hybrid visualization presupposes a different approach: AI functions as an instrument in the hands of a cognitive architect. In this case, authorship is assigned to the human if they make a substantial contribution to the formation of the training dataset, the choice of neural network parameters, and the final curatorship. Table 12 presents the legal matrix of responsibility and authorship in hybrid visualization.

Table 12. Legal matrix of responsibility and authorship in hybrid visualization (Quintais, 2025)

Intellectual property object	Legal status	Governing act / precedent	Requirements for protection
Neural profile (Personal AI Profile)	Complex software–creative object	USCO Part 2 (2025), EU AI Act	Training on exclusive, author-owned data; evidence of creative selection/curation of references.
AI-processed photo (AI-Assisted Work)	Protected by the human author's copyright	Beijing Internet Court decision (2023)	Proof of substantial human control over the final expressive outcome.
Training dataset (Input side)	Subject to licensing or TDM exceptions	CDSM Directive (Art. 4), USCO Part 3 (2025)	Compliance with the opt-out mechanism; no pirated content in the dataset.
Style (Artistic Style)	Generally, an unregulated idea	ArtBulb Framework, AICD Dataset	Recognition of style as expression when there is a unique combination of visual features.

The European Union, through the AI Act (EU AI Act 2024/1689), introduces strict transparency requirements for general-purpose AI models (GPAI). Users are obliged to disclose the fact of AI use if the result has been substantially altered by algorithms (Quintais, 2025). In commercial photography, this leads to the implementation of content provenance systems, such as the C2PA standard, which embeds the history of all modifications in file metadata, thereby confirming the image's authenticity (Quintais, 2025).

The ethical dimension concerns algorithmic bias. In 2024–2025, studies confirmed the tendency of neural networks to distort skin tones of certain population groups due to imbalances in training datasets (Alipour et al., 2024). The use of the Monk Skin Tone (MST) scale, with 10 types, rather than the outdated 6-type Fitzpatrick scale, is an ethical imperative for ensuring inclusivity in commercial photography.

To minimize legal risks within the hybrid methodology, the following algorithm of actions is recommended.

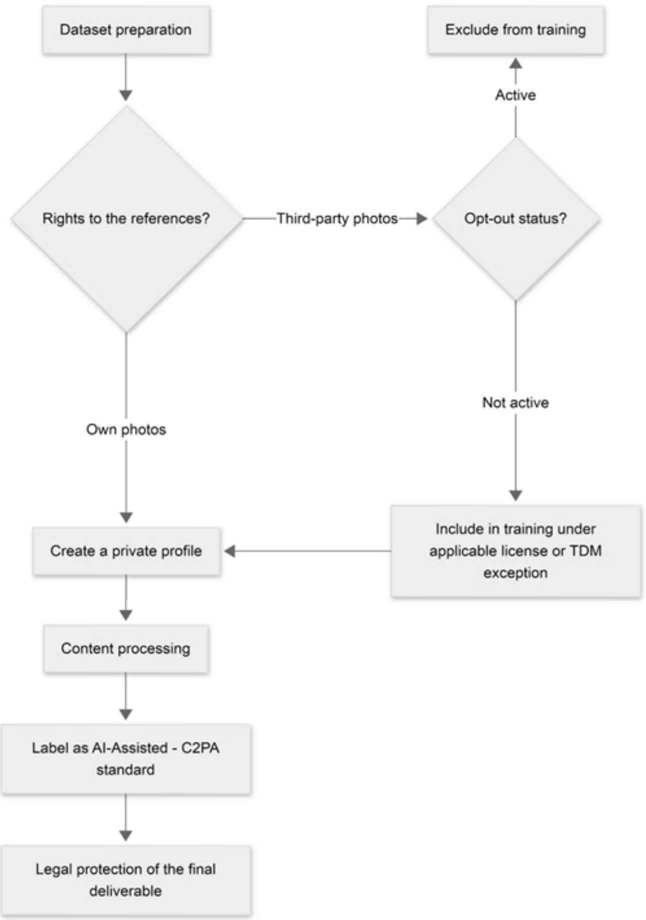

Fig. 18. Algorithm for legal support of a hybrid work process

The algorithm formalizes the path of legitimacy for a pipeline in which generative processing relies on references: it resolves the question of data origin first and only thereafter permits training and output. The key idea is to temporally separate two types of risk: the risk of unlawful use of input materials (references) and the risk of incorrect attribution/transparency at the output (labeling). Such a decomposition reduces process entropy: legal uncertainties are embedded into early branches, so downstream stages (profile creation, content processing) become more deterministic and reproducible.

The first branch sets the mode of legal control: for one's own photographs, the system almost immediately creates a private profile, thereby increasing the demonstrability of authorial contribution and reducing the likelihood of conflicts of rights. For third-party references, an opt-out check is triggered, functioning as a binary access gate. If opt-out is active, the data are excluded from training. If it is not active, the algorithm allows inclusion but requires a legal basis (a license or a TDM exception). Otherwise, the protocol's logic becomes internally contradictory. An important nuance is that including the training node effectively requires an audit of sources and metadata. Without this, the condition not active does not guarantee that the dataset is permissible (for example, pirated content or lack of a valid source violates the underlying assumptions).

The output part (processing, C2PA labeling, legal protection) operates as a traceability loop: labeling records the fact of AI-assisted intervention and provides cryptographically convenient evidence of content provenance, which can subsequently support rights claims and reduce the risk of misleading product presentation. At the same time, the algorithm constructs an evidentiary structure: it minimizes the space of contentious interpretations but does not eliminate it entirely, because the criteria of substantial human control and the boundaries of TDM/licensing remain contextual. In terms of scientific methodology, this is akin to experimental design: one does not prove the hypothesis in advance, but structures the protocol so that, upon testing, there are fewer hidden variables and more verifiable traces.

Particular attention is paid to the issue of training AI on the works of living authors. In 2025, the U.S. Copyright Office, in the third part of its report, emphasizes that deliberate training of models on pirated datasets almost certainly falls outside the bounds of fair use (U.S. Copyright Office, 2025). This creates demand for remuneration schemes for photographers whose work is used as references for global neural networks.

Thus, Chapter 4 demonstrates that practical implementation of the hybrid visualization methodology is possible only under the condition of strict quality standardization, understanding of scaling mechanisms, and adherence to ethical and legal frameworks. This transforms the technology from a set of disparate tools into a coherent industrial system prepared for the challenges of the knowledge economy.

CONCLUSION

The study has shown that contemporary commercial photography has simultaneously reached an aesthetic plateau of digital sterility and an economic limit of manual post-production, where selection and color correction become key bottlenecks, amplifying decision fatigue and burnout risk. Under these conditions, algorithmization of the creative process is justified as a response to the conflict between mass content production and the requirement for its emotional and human expressiveness, which the market increasingly associates with the Film Look.

The key theoretical result lies in explaining why analog aesthetics are perceived as premium psychophysiologically: film grain is interpreted by the visual system as a living stochastic texture, and the nonlinear subtractive color logic and soft compression of highlights (highlight roll-off) provide more perceptually comfortable transitions compared to the linear digital model and harsh clipping. Thus, filmness in the work is formalized as a set of measurable features (tonal curve toe–gamma–shoulder, texture structure, behavior at extreme luminance levels), rather than as a subjective stylization.

The principal methodological result is the development of an authorial technology of hybrid visualization, in which digitized film acts as ground truth for calibrating neural profiles. It is shown that the transition from static presets to neural profiles requires a rigorously organized dataset of digital RAW + film scan pairs with control of scanner domain shifts and formalized validation: keypoint detection (ORB/SIFT), geometric alignment (RANSAC + homography) with a reprojection error < 0.5 px, and luminance normalization via histogram matching in the CIELAB L-channel. Training is described as the search for a transformation function implemented through a Neural 3D LUT and optimized by a combined loss function, where priority is given to the accuracy of film color and perceptual structure, and texture is modeled as luminance-dependent grain, i.e., as a parameterized stochastic process rather than uniformly overlaid noise.

A distinct key result is the introduction of the Iterative Model Calibration (IMC) loop, which translates model improvement into a regime of reproducible engineering control: after each epoch, quality is diagnosed using ΔE-ITP, difficult scenes are singled out as risk zones, fine-tuning is intensified precisely on problematic segments, and final stability is confirmed by independent validation. Within this same logic, the most vulnerable area of commercial practice, skin tones, is formalized: biometric bias and typical shifts are decomposed along the MST scale (10 types) and the CIELAB/ITA parameters, making corrections more targeted and verifiable rather than taste-based.

The practical result of the monograph is the reconstruction of the full production cycle as a hybrid cyber-physical system: film is assigned to key shots (approximately one quarter of the final selection) and carries the function of a coloristic anchor, while digital serves high entropy, dynamics, mixed light, and low light as a circuit of controllability and insurance. Post-production is translated into a pipeline consisting of AI culling (reducing the set by 70–80%), batch neural processing with film matching, and a limited Human Touch stage (5–10% of operations), where the human remains the final arbiter of meaning, ethics, and serial integrity. The quantitative efficiency model (TDABC) on a typical wedding project demonstrates a reduction of specialist involvement time from approximately 27 to 3 hours (−89%) and a decrease in overall cost, even accounting for film and computation expenses, thereby transforming the premium Film Look from an expensive embellishment into an economically sustainable scaling strategy.

Quality standardization in the work is framed as a transition from diffuse aesthetic judgments to a system of separate measuring instruments: technical quality is used for rejecting defective frames, aesthetic value for ranking commercial appeal, perceptual correspondence to film (EITP, LPIPS) for profile calibration, and stylistic unity for controlling series integrity. An important conclusion here is that the quality of the hybrid result must be confirmed not by the perfection of a single frame but by the stability of the error distribution and the absence of artifacts noticeable in serial viewing.

The legal and ethical component elevates the technology to the level of implementable industrial practice: the status of the result is fixed as AI-assisted work in the presence of substantial human control, and the neural profile is construed as a complex software–creative object, provided that it is trained on proprietary data with demonstrable curatorial selection. To reduce risks, principles of transparency and traceability are proposed (including content-provenance labeling via C2PA), as well as the requirement to exclude pirated datasets and to respect opt-out mechanisms, since it is precisely the data's input legitimacy that determines the robustness of the entire circuit. The ethical minimum in commercial work is defined through mandatory control of skin tones using MST and metrics sensitive to group-specific shifts.

Implementation recommendations converge on procedural discipline, without which hybrid visualization degrades into a set of disjointed effects. On the data side, it is critical to construct paired RAW–scan datasets with strict geometric and luminance validation and to preselect which scanner signature will serve as the reference, to avoid introducing uncontrolled bias into the model. On the training side, IMC should be maintained as a mandatory operational regime for the profile, measuring progress via ΔE-ITP and separately monitoring skin tones via MST/ITA, while on the production side, the protocol of medium allocation, time synchronization, and the chain of film custody should be fixed in a Chain-of-Custody-like manner so that the physical circuit does not disrupt digital traceability. In post-production, the boundary of automation should be preserved: routine is delegated to algorithms, and the Human Touch is retained as a limited but mandatory stage of artistic and ethical validation of the series.

The prospects for development arising from the monograph include further scalability through agent-based and cloud infrastructures, where processing stages become services with scenario-based control, and studios transition to protected environments for deploying their own models (Sovereign AI) for privacy and strategic independence. In applied terms, this implies an increase in personalization accuracy, up to digital twins of an authorial style, while simultaneously strengthening requirements for the standardization of quality metrics, for transparency of content provenance, and for raising AI fluency among practitioners as a key condition for overcoming the skills deficit barrier in the implementation of algorithmized workflows.

REFERENCES

1. Adobe Exchange. (2022, November 25). *Neurapix AI Editing for Lightroom Classic*. Adobe Exchange. https://exchange.adobe.com/apps/cc/109033/neurapix-smartpresets-for-lightroom-classic

2. Afifi, M., Brubaker, M. A., & Brown, M. S. (2021). Auto White-Balance Correction for Mixed-Illuminant Scenes. *ArXiv*. https://doi.org/10.48550/arXiv.2109.08750
3. Alipour, N., Burke, T., & Courtney, J. (2024). Skin Type Diversity in Skin Lesion Datasets: A Review. *Current Dermatology Reports*, *13*, 198–210. https://doi.org/10.1007/s13671-024-00440-0
4. AMI. (2024, August 19). *Burnout hits 70% of media, marketing and creative professionals, reveals 2024 Mentally Healthy Survey*. Australian Marketing Institute. https://ami.org.au/knowledge-hub/burnout-hits-70-of-media-marketing-and-creative-professionals-reveals-2024-mentally-healthy-survey/
5. Audiffren, M., Lefebvre, R., Silvestrini, N., Fisher, J., Ravel, S., & Pageaux, B. (2023). Editorial: Effort-based decision-making and cognitive fatigue. *Frontiers in Neuroscience*, *17*. https://doi.org/10.3389/fnins.2023.1230022
6. Azuddin, A. (2024, October 15). *Margot Robbie and Jacob Elordi Star in New Chanel N°5 Short Film*. Harper's BAZAAR. https://harpersbazaar.my/beauty/see-you-at-5-chanel/
7. Babashahi, L., Barbosa, C. E., Lima, Y., Lyra, A., Salazar, H., Argôlo, M., Almeida, M. A. de, & Souza, J. M. de. (2024). AI in the Workplace: a Systematic Review of Skill Transformation in the Industry. *Administrative Sciences*, *14*(6), 127. https://doi.org/10.3390/admsci14060127
8. Bervar, M., Bertoncel, T., & Bach, P. (2026). Generative Artificial Intelligence and the Creative Industries: A Bibliometric Review and Research Agenda. *Systems*, *14*(2), 138. https://doi.org/10.3390/systems14020138
9. BLS. (2023, April 25). *27-4021 Photographers*. Bureau of Labor Statistics. https://www.bls.gov/oes/2023/may/oes274021.htm
10. Bray, I., Lerigo-Sampson, M., Morey, Y., & Williams, J. (2024). Mental health of social media influencers. *Journal of Occupational Health*, *66*(1). https://doi.org/10.1093/joccuh/uiae045
11. Clarke, G. (2025, March 29). *Abby Champion Stars in Tiffany & Co.'s New High Jewelry Campaign*. Lofficielusa. https://www.lofficielusa.com/fashion/abby-champion-tiffany-co-high-jewelry-campaign-images
12. Conde, M. V., Vazquez-Corral, J., Brown, M. S., & Timofte, R. (2024). NILUT: Conditional Neural Implicit 3D Lookup Tables for Image Enhancement. *Proceedings of the AAAI Conference on Artificial Intelligence*, *38*(2), 1371–1379. https://doi.org/10.1609/aaai.v38i2.27901
13. Darmawan, N. Y. S., Amir, Y., Iwan, A., & Mutiaz, I. R. (2023). Analog Photography in the Digital Age: Examining Transformation, Alienation and Authenticity in Modern Photographic Practice. *Indonesian Journal of Art and Design Studies*, *2*(3), 217–232. https://doi.org/10.55927/ijads.v2i3.11019
14. Di Natale, A. F., Simonetti, M. E., La Rocca, S., & Bricolo, E. (2023). Uncanny valley effect: A qualitative synthesis of empirical research to assess the suitability of using virtual faces in psychological research. *Computers in Human Behavior Reports*, *10*, 100288. https://doi.org/10.1016/j.chbr.2023.100288
15. Drassinower, A. (2022). Remarks on Technological Neutrality in Copyright Law as a Subject Matter Problem: Lessons From Canada. *The Cambridge Law Journal*, *81*(1), 50–83. https://doi.org/10.1017/s0008197321001045
16. Gong, H., Su, J., Seng, K. P., Nguyen, A., Liu, A., & Liu, H. (2023). Film-GAN: towards realistic analog film photo generation. *Neural Computing and Applications*, *36*(8), 4281–4291. https://doi.org/10.1007/s00521-023-09283-5
17. Grigonis, H. K. (2025, December 19). *Photographers saved 89 million hours – 12 work weeks each – using AI in 2025, study suggests*. Digital Camera World. https://www.digitalcameraworld.com/tech/artificial-intelligence/photographers-saved-89-million-hours-12-work-weeks-each-using-ai-in-2025-study-suggests
18. Guo, W., Zhang, W., Zhang, J., Zhang, S., & Jin, L. (2024). The Application of AI Technology in Post-Processing of Photography. *Proceedings of the 3rd International Conference on Educational Innovation and Multimedia Technology*. https://doi.org/10.4108/eai.29-3-2024.2347637
19. Guo, X., Li, Z., Tian, D., Zhang, K., Pappas, T. N., Egmond, R. van, & Ridder, H. de. (2024). Analysis of User Preferences for Film Grain Noise in Images and Video Sequences. *Electronic Imaging*, *36*(11), 230. https://doi.org/10.2352/ei.2024.36.11.hvei-230
20. Haim, I. (2025, October 22). *Aftershoot vs. Imagen: A Photographer's In-Depth Comparison*. Imagen AI. https://imagen-ai.com/valuable-tips/aftershoot-vs-imagen/
21. Haim, I. (2026, January 19). *10 Best AI Photo Editor in 2026: Reviews & Buying Guide*. Imagen AI. https://imagen-ai.com/valuable-tips/best-ai-photo-editor-guide/
22. Hausken, L. (2024). Photorealism versus photography. AI-generated depiction in the age of visual disinformation. *Journal of Aesthetics & Culture*, *16*(1), 2340787. https://doi.org/10.1080/20004214.2024.2340787
23. Jpneylon. (n.d.). *Discuss the characteristics of the HD curve*. Jpneylon. Retrieved February 5, 2026, from https://jpneylon.github.io/ABR/ABR_Study/1-Equipment/16-Film.pdf
24. Karimipour, H., & Witzel, C. (2024). Colour expectations

across illumination changes. *Vision Research*, *222*, 108451. https://doi.org/10.1016/j.visres.2024.108451

25. Kirchner, E., van Wijk, C., van Beek, H., & Koster, T. (2021). Exploring the limits of color accuracy in technical photography. *Heritage Science*, *9*, 57. https://doi.org/10.1186/s40494-021-00536-x
26. Kislinger, L., & Kotrschal, K. (2021). Hunters and Gatherers of Pictures: Why Photography Has Become a Human Universal. *Frontiers in Psychology*, *12*(12). https://doi.org/10.3389/fpsyg.2021.654474
27. Mackenzie, P., Senghaas, M., & Achddou, R. (2024). CNNs for Style Transfer of Digital to Film Photography. *ArXiv*. https://doi.org/10.48550/arXiv.2411.15967
28. Margadona, L. A. (2023). Understanding the New Analogue Photography: Meanings, Methodologies and Aesthetics. *New Explorations*, *3*(1). https://doi.org/10.7202/1097584ar
29. Mayer, H., Yee, L., Chui, M., & Roberts, R. (2025, January 28). *Superagency in the workplace: Empowering people to unlock AI's full potential*. McKinsey & Company. https://www.mckinsey.com/capabilities/tech-and-ai/our-insights/superagency-in-the-workplace-empowering-people-to-unlock-ais-full-potential-at-work
30. Melzi, P., Rathgeb, C., Tolosana, R., Vera-Rodriguez, R., Morales, A., Lawatsch, D., Domin, F., & Schaubert, M. (2024). Synthetic Data for the Mitigation of Demographic Biases in Face Recognition. *ArXiv*. https://doi.org/10.48550/arXiv.2402.01472
31. Mordor Intelligence. (2025). *North America Photographic Services Market Size & Share Analysis*. Mordor Intelligence. https://www.mordorintelligence.com/industry-reports/north-america-photographic-services-market
32. O'Brien, K., Katrak, F., & Lartigue, J.-P. (2025, February 10). *Winning at digital transformation and cloud through Hybrid by Design*. IBM. https://www.ibm.com/think/insights/winning-at-digital-transformation-and-the-cloud-through-hybrid-by-design
33. Okada, K., & Motoyoshi, I. (2021). Human Texture Vision as Multi-Order Spectral Analysis. *Frontiers in Computational Neuroscience*, *15*, 692334. https://doi.org/10.3389/fncom.2021.692334
34. Osińska, V., Kortas, W., Szalach, A., & Welter, M. (2025). AI Images vs. Real Photographs: Investigating Visual Recognition and Perception. *Journal of Eye Movement Research*, *18*(6), 61. https://doi.org/10.3390/jemr18060061
35. Öztaş, Y. E., & Arda, B. (2025). Re-evaluating creative labor in the age of artificial intelligence: a qualitative case study of creative workers' perspectives on technological transformation in creative industries. *AI & Society*, *40*, 4119–4130. https://doi.org/10.1007/s00146-025-02180-6
36. Petrović, S. R. (2024, October 14). *CHANEL N°5 Campaign With Margot Robbie: See You At 5!* Fragrantica. https://www.fragrantica.com/news/CHANEL-N05-Campaign-With-Margot-Robbie-See-You-At-5-21259.html
37. PhotoVision. (n.d.). *Noritsu vs. Frontier*. PhotoVision. Retrieved January 7, 2026, from https://www.photovision.co/working-with-us/film-scannerbattle/
38. Plutino, A., Armellin, L., Sarti, B., & Rizzi, A. (2025). Tonal distribution and glare assessment in cinematographic film scanners. *Signal Image and Video Processing*, *19*(12). https://doi.org/10.1007/s11760-025-04569-8
39. Quintais, J. P. (2025). Generative AI, copyright and the AI Act. *Computer Law & Security Review*, *56*, 106107. https://doi.org/10.1016/j.clsr.2025.106107
40. Richard Photo Lab. (2023, April 25). *It's back... Noritsu vs Frontier!* Richard Photo Lab. https://richardphotolab.com/blogs/post/its-back-noritsu-vs-frontier
41. Romanov, O. Y. (2024). The hybridization of visual culture: The revival of analog photography in the digital age. *Manuscript*, *17*(4), 673–677. https://doi.org/10.30853/mns20240100
42. Seeram, E. (2023). Digital Image Quality Descriptors and Performance Characteristics. In *X-Ray Imaging Systems for Biomedical Engineering Technology* (pp. 67–76). Springer Nature. https://doi.org/10.1007/978-3-031-46266-5_6
43. Shah, P., Sankhe, D., Rashid, M., Khaled, Z., Puyol-Antón, E., Lee, T., Alqarni, M., Rai, S., & King, A. P. (2025). The Impact of Skin Tone Label Granularity on the Performance and Fairness of AI Based Dermatology Image Classification Models. *ArXiv*. https://doi.org/10.48550/arxiv.2509.11184
44. Shakya, S., Mudiyanselage, S. B., Robinson, S., Randall, S., & Gao, L. (2025). Time-Driven Activity-Based Costing and Its Use in Health Economic Analysis: A Systematic Literature Review. *Applied Health Economics and Health Policy*, *23*, 989–1004. https://doi.org/10.1007/s40258-025-00988-3
45. Shi, Q., & Shen, L. (2025). Advancing Firm-Level Digital Technology Diffusion: A Hybrid Bibliometric and Framework-Based Systematic Literature Review. *Systems*, *13*(4), 262. https://doi.org/10.3390/systems13040262
46. Silva Jasaui, D., Martí-Testón, A., Muñoz, A., Moriniello, F., Solanes, J. E., & Gracia, L. (2024). Virtual Production: Real-Time Rendering Pipelines for Indie Studios and the Potential in Different Scenarios. *Applied Sciences*, *14*(6), 2530. https://doi.org/10.3390/app14062530

47. Singla, A., Sukharevsky, A., Hall, B., Yee, L., Chui, M., & Balakrishnan, T. (2025, March 12). *The state of AI: How organizations are rewiring to capture value*. McKinsey & Company. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai

48. Sun, L. (2024). Aesthetic Trends and Evolutions in Photography Art in the Digital Era. *Academic Journal of Management and Social Sciences*, *7*(3), 107–110. https://doi.org/10.54097/whhc5x16

49. Tam, F. Y., & Lung, J. (2024). Digital marketing strategies for luxury fashion brands: A systematic literature review. *International Journal of Information Management Data Insights*, *5*(1), 100309. https://doi.org/10.1016/j.jjimei.2024.100309

50. Taylor, J., & Xu, Y. (2022). Representation of Color, Form, and their Conjunction across the Human Ventral Visual Pathway. *NeuroImage*, *251*, 118941. https://doi.org/10.1016/j.neuroimage.2022.118941

51. Trigka, M., & Dritsas, E. (2025). A Comprehensive Survey of Deep Learning Approaches in Image Processing. *Sensors*, *25*(2), 531. https://doi.org/10.3390/s25020531

52. Tsiavos, V., & Kitsios, F. (2025). The digital transformation of the film industry: How Artificial Intelligence is changing the seventh art. *Telecommunications Policy*, *49*(8), 103021. https://doi.org/10.1016/j.telpol.2025.103021

53. U.S. Copyright Office. (2025). *Copyright and Artificial Intelligence Part 2: Copyrightability*. U.S. Copyright Office. https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-2-Copyrightability-Report.pdf

54. Velásquez-Salamanca, D., Martín-Pascual, M. Á., & Andreu-Sánchez, C. (2025). Interpretation of AI-Generated vs. Human-Made Images. *Journal of Imaging*, *11*(7), 227. https://doi.org/10.3390/jimaging11070227

55. Watkins, J. (2024). Alleviating digital fatigue through embodied artistic practice and green space. *International Journal of Performance Arts and Digital Media*, *20*(3), 527–542. https://doi.org/10.1080/14794713.2024.2305448